

Is Big Data a Big Deal?

Not Without Correct
Software!

Gregory M. Kapfhammer

creative commons licensed (BY-NC-ND) flickr photo shared by sunface13

What is Big
Data?

What is Big Data?

Electronic music group from New York City?

What is Big Data?

Electronic music group from New York City?
Yes, but not in this context!

What is Big Data?

Electronic music group from New York City?

Yes, but not in this context!

Create **models**, make **predictions**, and take **action**!



"v" Times Three?

flickr photo shared by Elad R under a [Creative Commons \(BY-NC \)](#) license

"v" Times Three?

Volume

flickr photo shared by Elad R under a [Creative Commons \(BY-NC \)](#) license

"v" Times Three?

Volume
Velocity

flickr photo shared by Elad R under a [Creative Commons \(BY-NC \)](#) license

"v" Times Three?

Volume
Velocity
Variety

flickr photo shared by Elad R under a [Creative Commons \(BY-NC \)](#) license

Wait, have we
missed a "v"?





We Need Veracity!

flickr photo shared by [Brandon Giesbrecht](#) under a [Creative Commons \(BY \)](#) license



We Need Veracity!

Confidence in the **correctness** of the data and software

flickr photo shared by **Brandon Giesbrecht** under a **Creative Commons (BY)** license



We Need Veracity!

Confidence in the **correctness** of the data and software
Achieved through **verification** and **validation**

flickr photo shared by **Brandon Giesbrecht** under a **Creative Commons (BY)** license

Software Testing

Software Testing

Execute the test cases against the system under test

Software Testing

Execute the test cases against the system under test
Check to see if the system's output is as expected



Status Report

Stocksnap.io photo shared by Olu Eletu under a Public Domain license

Status Report

Testing **data mining** and **database systems**

Stocksnap.io photo shared by Olu Eletu under a Public Domain license

Status Report

Testing **data mining** and **database systems**
Research is preliminary and transfer is not common!

[Stocksnap.io](#) photo shared by [Olu Eletu](#) under a [Public Domain](#) license

Challenges?

Challenges?

Testing: scale to the level of big data systems

Challenges?

Testing: scale to the level of big data systems

Big Data: Holistically consider all of the four "V"s

Next Steps?



A photograph of two people sitting at a wooden table in a cafe. A dark semi-transparent overlay covers the center of the image, containing the text 'Let's Talk!'. Below the text, there is a line of smaller text: 'Stocksnap.io photo shared by Alejandro Escamilla under a Public Domain license'. The background shows a brick wall, a brown leather bag, a glass of water, a smartphone, a tablet, and a cup of coffee.

Let's Talk!

Stocksnap.io photo shared by Alejandro Escamilla under a Public Domain license

Let's Talk!

Collaboration between research communities

Stocksnap.io photo shared by **Alejandro Escamilla** under a **Public Domain** license

Let's Talk!

Collaboration between research communities
Demonstrate the **correctness** of big data systems

Stocksnap.io photo shared by [Alejandro Escamilla](#) under a [Public Domain license](#)

Is Big Data a
Big Deal?

Of course it is!



A photograph of a group of children in a classroom, all wearing yellow polo shirts. They are cheering and clapping enthusiastically. A semi-transparent dark grey rectangle is overlaid in the center of the image, containing the title text. The background is slightly blurred, showing other children and a red wall.

A Bigger Deal

Stocksnap.io photo shared by [Aundre Larrow](#) under a [Public Domain](#) license

A group of diverse children in a classroom, some cheering and clapping, with a dark overlay containing text.

A Bigger Deal

Greater impact with **high-quality** big data systems

Stocksnap.io photo shared by **Aundre Larrow** under a **Public Domain** license

A vibrant, out-of-focus photograph of a diverse group of people, primarily children, cheering and clapping. The scene is bright and joyful, with people in the foreground and background all showing expressions of excitement. A semi-transparent dark grey rectangle is overlaid in the center, containing white and orange text.

A Bigger Deal

Greater impact with **high-quality** big data systems
Important societal benefits with **correct** predictions

Stocksnap.io photo shared by **Aundre Larrow** under a **Public Domain** license

A vibrant, out-of-focus background image showing a diverse group of people, including children and adults, cheering and clapping enthusiastically. The scene appears to be a public event or a school assembly. The text is overlaid on a dark, semi-transparent rectangular area in the center of the image.

A Bigger Deal

Greater impact with **high-quality** big data systems
Important societal benefits with **correct** predictions
Enhance and sustain our influence — **together!**

Stocksnap.io photo shared by **Aundre Larrow** under a **Public Domain** license