

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

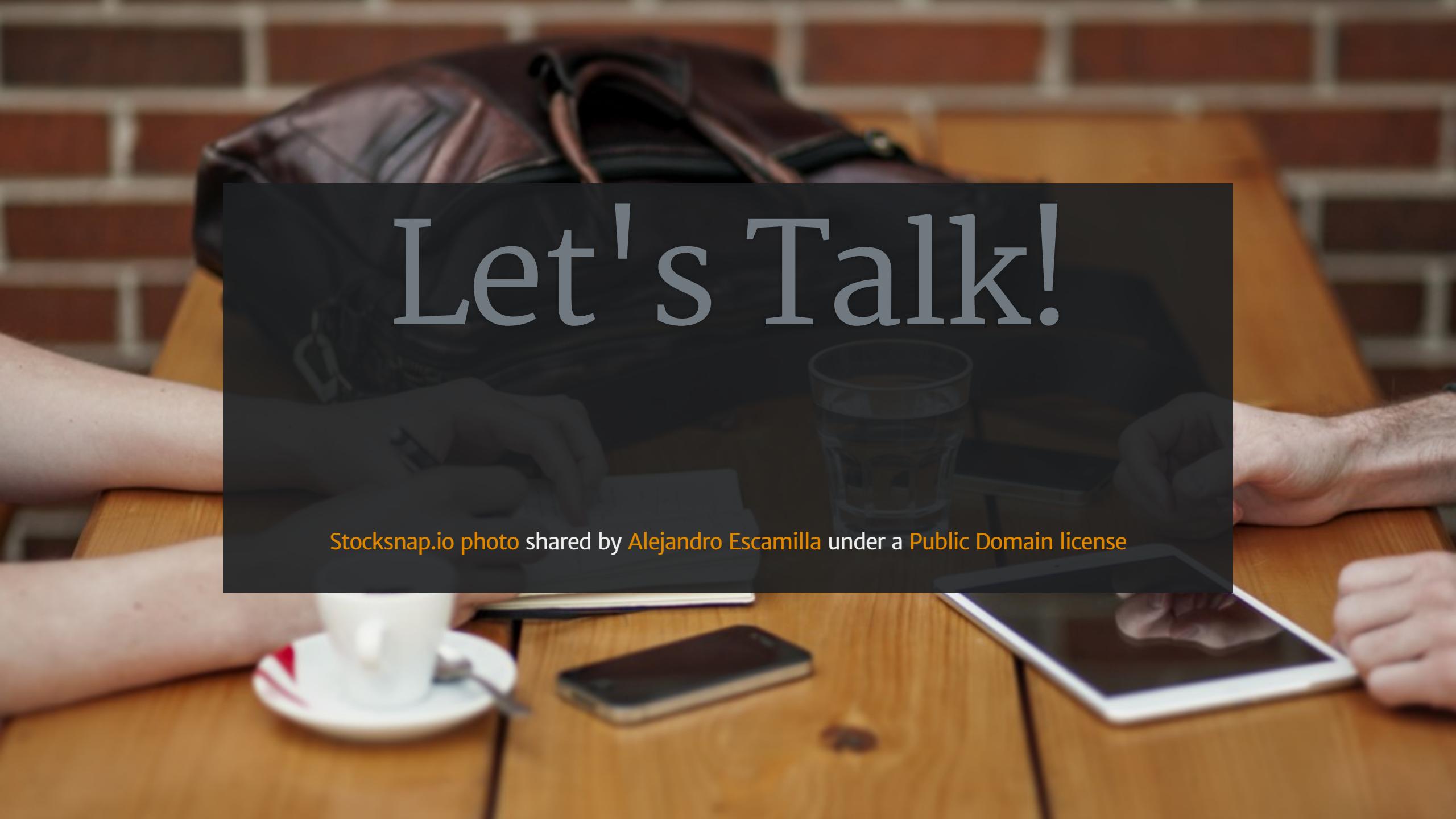
Challenges?

Testing: scale to the level of big data systems

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

Next Steps?



A photograph of a person sitting at a wooden table, viewed from the side. They are wearing a brown backpack and have their hands resting on the table. On the table are a white cup with a saucer, a smartphone, a tablet, a notebook, and a clear glass. In the background, there is a brick wall.

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 Bigger Deal

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

A Bigger Deal

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

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

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 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