

# Snowflake and the Data Marketplace

**Atlanta Snowflake Cloud Data Warehouse User Group**  
**Wednesday | December 11, 2019**  
**6:45 - 8:00 PM**



**Stewart Bryson, Red Pill Analytics**  
**@stewartbryson**



# Founder & CEO Red Pill Analytics

twitter: @stewartbryson

medium: @stewartbryson

linkedin: stewartbryson

A wide-angle photograph of a mountain range under a clear blue sky. The mountains are covered in white snow, with bright sunlight reflecting off the peaks. In the foreground and middle ground, there are rugged, rocky slopes partially covered in snow. The lighting suggests either sunrise or sunset, casting a warm orange glow on the exposed rock faces and the tops of the mountains.

Your Data Warehouse is elastic.  
Shouldn't your team be?

ELASTIC DELIVERY

# Elastic Crew



## MULTIFUNCTIONAL TEAM

Each Crew has the necessary players to move from question to insight in a single sprint.

## AUTONOMOUS

Whether you have an existing team or not, a Crew works autonomously to solve whatever data problems you assign to it.

## SCALABLE

Scale to answer all of your data requirements by adding additional Crews.

# Elastic Surge



## SUPPLEMENT YOUR TEAM

Increase your delivery capacity without adding fulltime resources.

## SEMI-PERMANENT

Designed as a long-term enhancement to organizations with functioning teams.

## RESPOND TO DEMAND

Be more responsive to user demand: increase in size, swap out functional roles, or simply add a new technical skill for a few sprints.

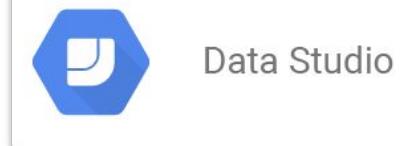
# Technologies We Use

## Data Warehouse



Amazon Athena

## Analytics



**ORACLE**  
ANALYTICS®

## Data-engineering & ETL





We're hiring.

@redpilla

@stewartbryson



# History Lesson





**Resistance is futile**

A photograph of a person standing on a large, rugged rock formation. The person is seen from behind, with their arms wide open, suggesting a sense of accomplishment or freedom. The sky is filled with dramatic, colorful clouds, transitioning from deep blues and purples to bright orange and yellow near the horizon. In the background, there are more hills and mountains under the vast sky.

# What's *changed*?

A close-up photograph of a man in a dark suit jacket, white shirt, and patterned tie. He is adjusting his tie with his right hand. The lighting is dramatic, highlighting the contours of his neck and the texture of his suit. The background is dark and out of focus.

Not all data is *corporate* data.

```
CREATE TABLE test ( a INTEGER, b TEXT, c
```

```
SELECT * FROM
```



Not all data is *relational* data.



...and the *Cloud*



**Every 3 minutes** someone  
is diagnosed **with a blood cancer.**

**360 people each day**  
**15 people each hour**

**YOU  
COULD BE  
THE CURE**





# Online Registration

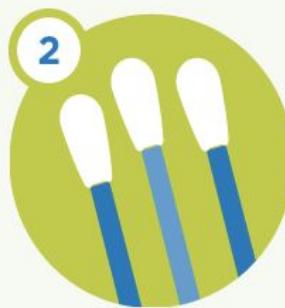
Have Questions?



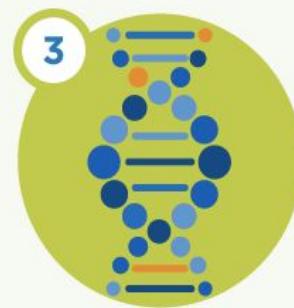
(800) MARROW-2



REGISTER  
ONLINE



RECEIVE & RETURN  
CHEEK SWAB KIT



BE READY IF CALLED  
AS A MATCH

# HapLogic: A Predictive Human Leukocyte Antigen-Matching Algorithm to Enhance Rapid Identification of the Optimal Unrelated Hematopoietic Stem Cell Sources for Transplantation

Jason Dehn <sup>1,\*</sup>, Michelle Setterholm <sup>1</sup>, Kelly Buck <sup>1</sup>, Jane Kempenich <sup>1</sup>, Beth Beduhn <sup>1</sup>, Loren Gragert <sup>1,2</sup>, Abeer Madbouly <sup>1</sup>, Stephanie Fingerson <sup>1</sup>, Martin Maiers <sup>1</sup>

<sup>1</sup> National Marrow Donor Program, Minneapolis, Minnesota

<sup>2</sup> Department of Pathology and Laboratory Medicine, Tulane Cancer Center, Tulane University School of Medicine, New Orleans, Louisiana

---

## Article history:

Received 10 May 2016

Accepted 29 July 2016

---

## Key Words:

HLA matching

Predictive algorithm

Donor selection

Cord blood unit selection

## A B S T R A C T

The search for a suitable human leukocyte antigen (HLA)-matched unrelated adult stem cell donor (URD) or umbilical cord blood unit (UCB) is a complex process. The National Marrow Donor Program (NMDP) developed a search algorithm known as HapLogic, which is currently provided within the NMDP Traxis application. The HapLogic algorithm has been in use since 2006 and has advanced URD/UCB HLA-matching technology. The algorithm has been shown to have high predictive accuracy, which can streamline URD/UCB selection and drive efficiencies in the search process to the benefit of the stem cell transplantation community. Here, we describe the fundamental components of the NMDP matching algorithm, output, validation, and future directions.

© 2016 American Society for Blood and Marrow Transplantation.

---

## INTRODUCTION

The search for a suitable human leukocyte antigen (HLA)-matched unrelated adult stem cell donor (URD) or umbilical cord blood unit (UCB) is a complex process, influenced by

in the international pool of stem cell sources, with a long potential URD list and often a clear high-resolution-matched URD, whereas others are more unique in that no available URD represents a potential full HLA match. Patient HLA types with

# Objectives





Gain  
*efficiencies*  
from a  
modern data  
platform.



Lower ***total*** cost of ownership.



Promote  
*real* data  
governance.



Enable  
*discovery*  
and self  
service.

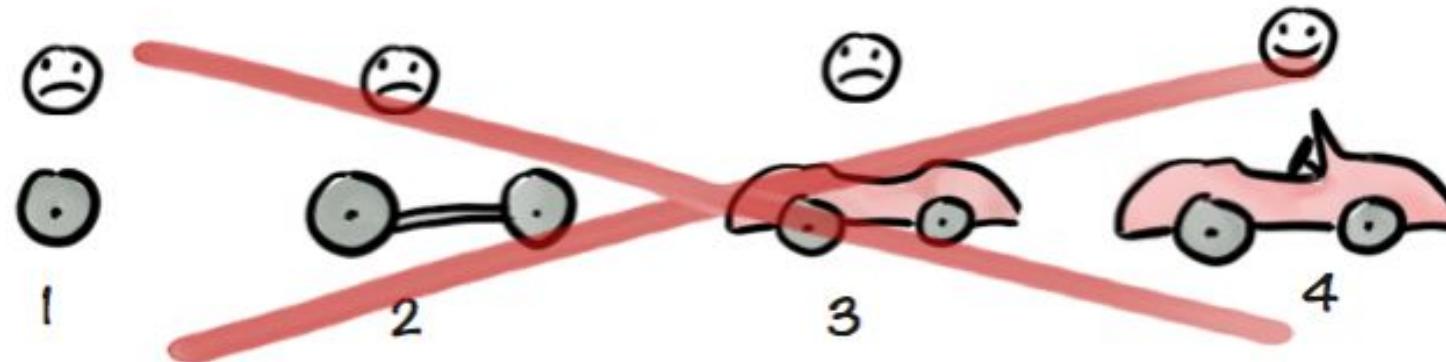




THINK BIG. start small. Move *fast*

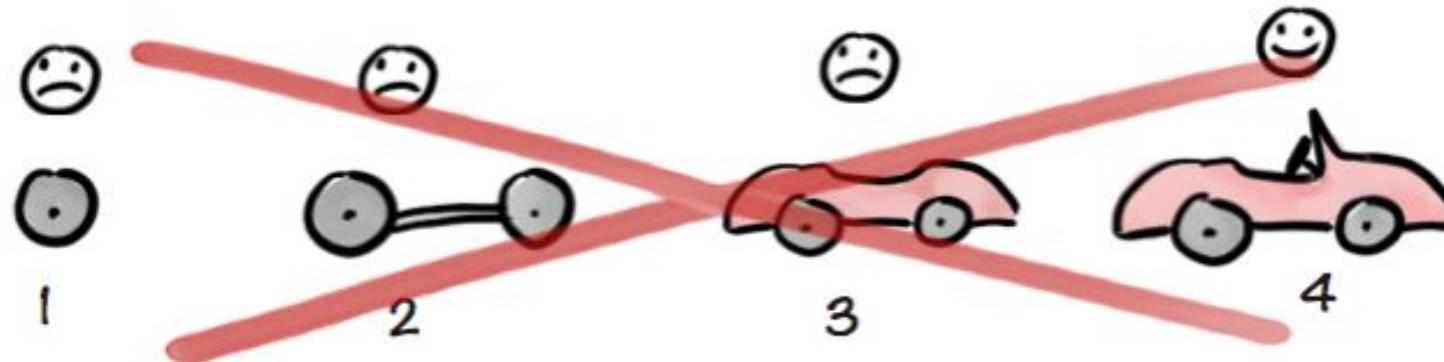
Not like this....

Image Source: Spotify

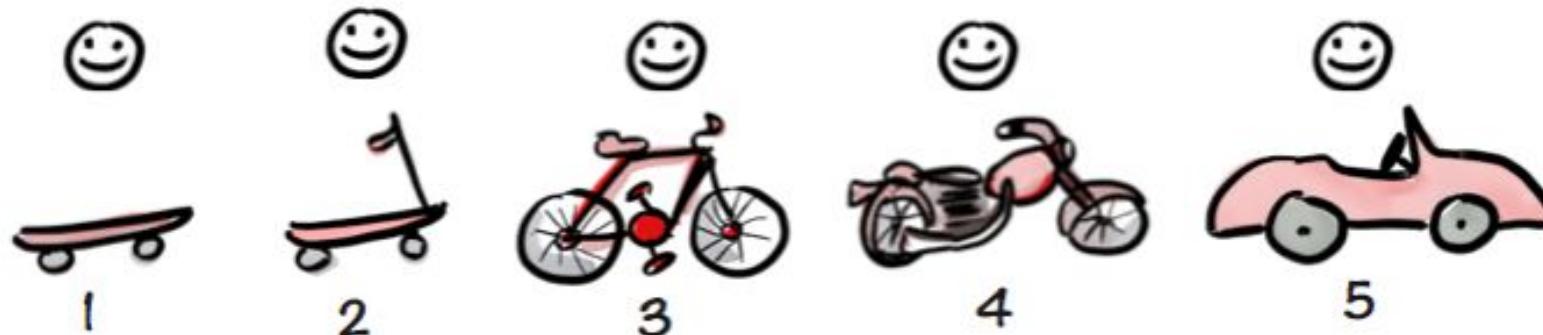


**Not like this....**

Image Source: Spotify



**Like this!**



Spend time on tasks that require  
*human* innovation.





Easy to Provision

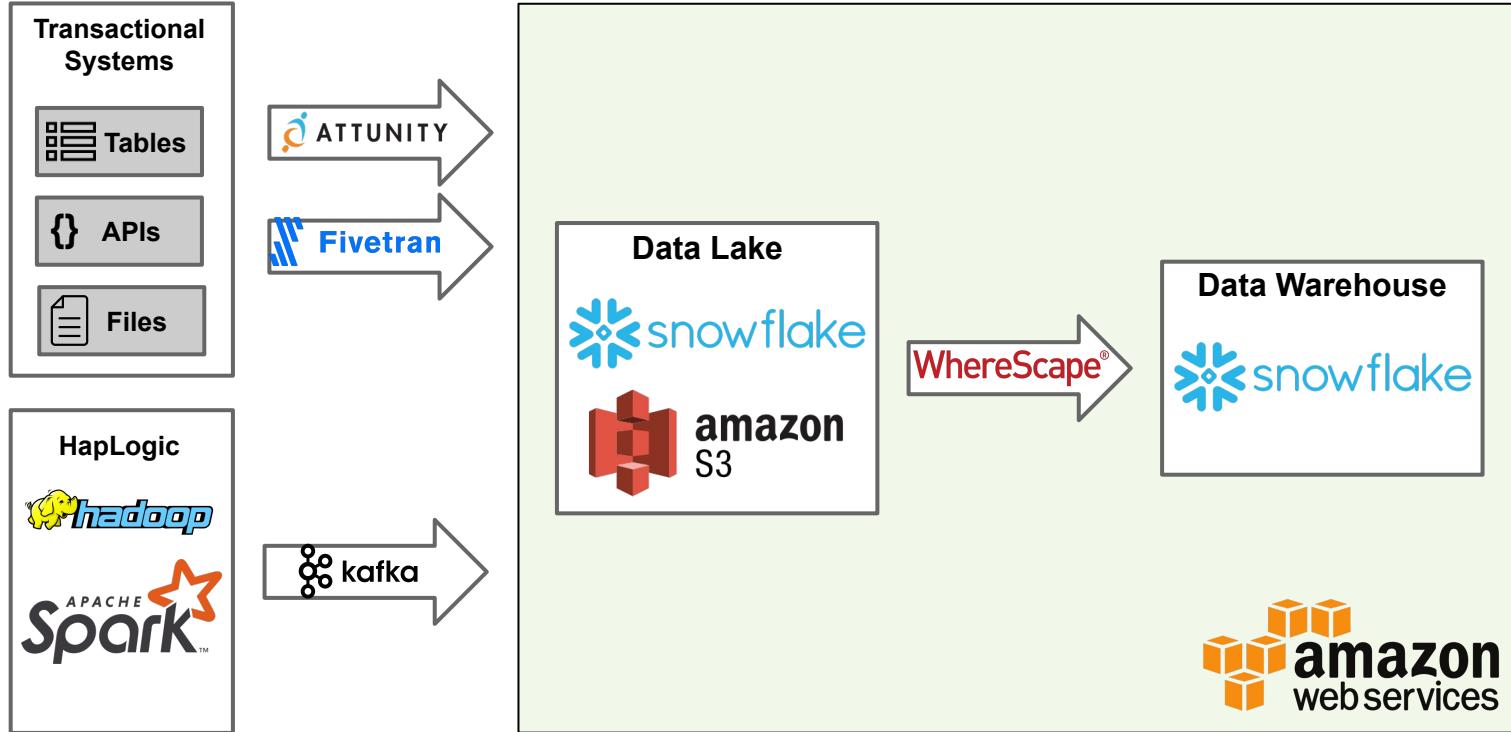


Easy to Manage



Easy to Share

# Standard Cloud-Native Architecture



# Data Lake *versus* Data Warehouse

Raw Data

Zero Transformation

Zero Opinion

Little Governance

Dimensional Model

Curated

Opinionated

Full Governance

# Have we solved *every* use case?



# Data Analyst

I look at raw data to determine  
how to build *better analytics*,  
but I wish it performed better.

# Data Scientist

I need to run *machine learning* algorithms against the raw data, but I'm not allowed to see PII.

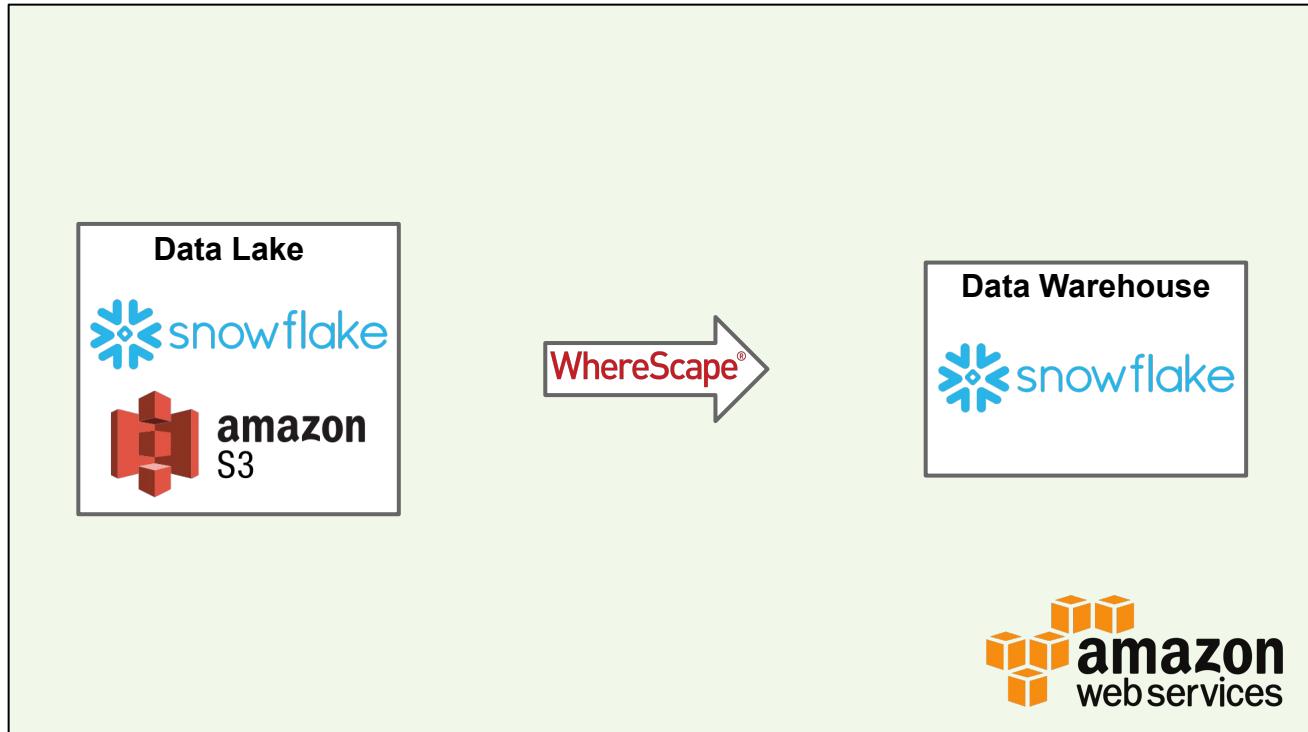
# Marketing Analyst

There's no marketing data in the data warehouse. I wish I could *make sense* of the data lake.

# Data Engineer

Instead of *delivering analytics*, I  
spend most of my time studying  
source system join patterns.

# Data Lake & Data Warehouse



# Enter the Data Marketplace



# The Data Marketplace

Common-sense transformations

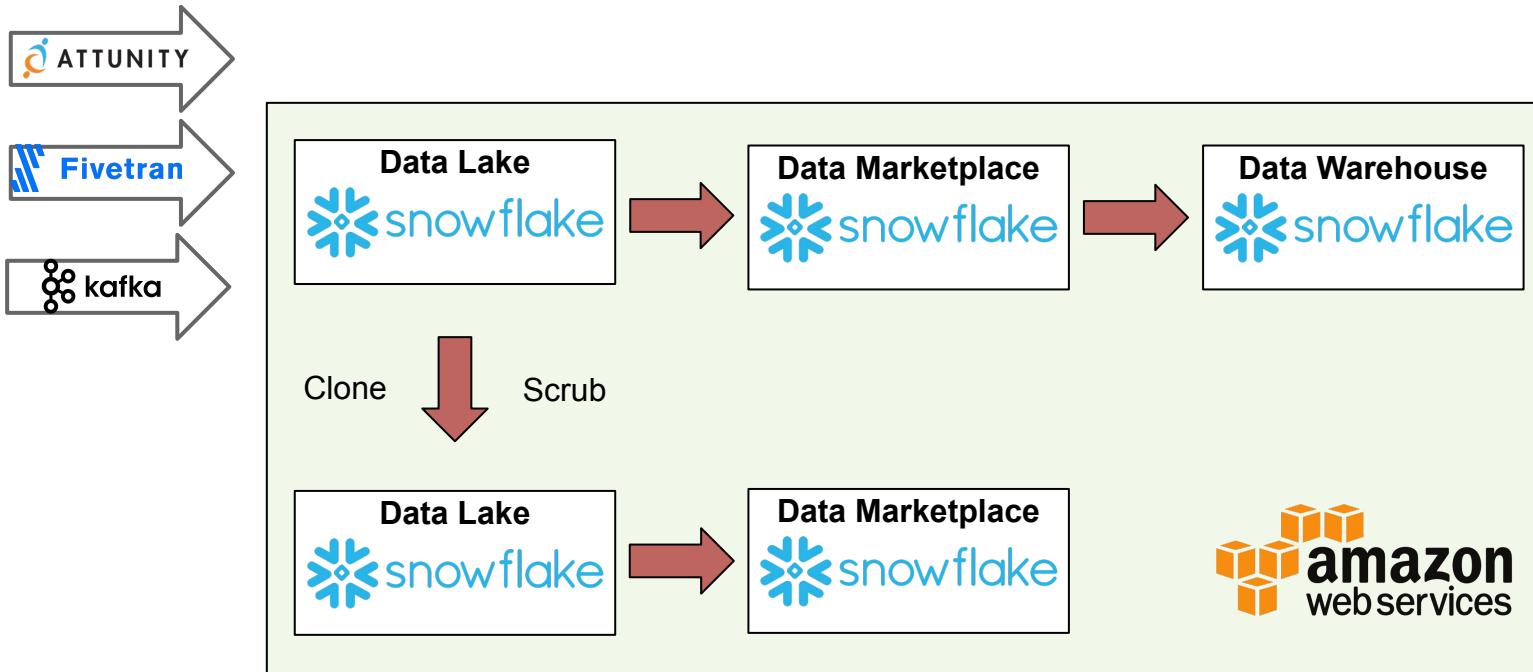
Full data governance

Passive data engineering  
services.

Proactive data engineering  
services



# Data Governance



looker



+ tableau®

ORACLE®  
ANALYTICS

The separation of *compute* and *storage* allows each department to bring its own budget.

# Triage Factors for Success



*questions?*

audience size  
& variety

answers?

*location of data*

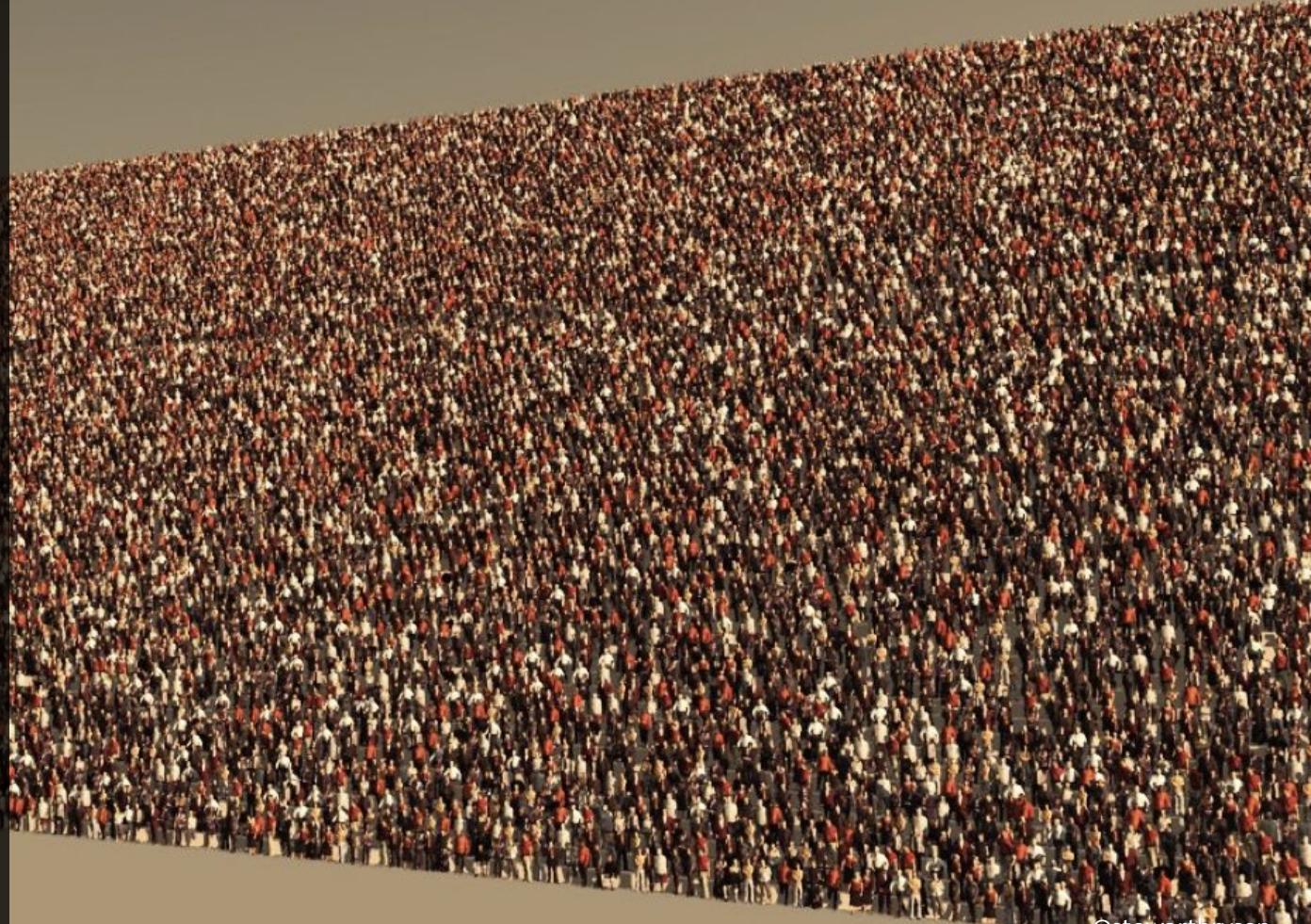
stability of  
solution

business users?

*power users?*

*frequency of data*

Audience  
*size* and  
*variety*



A horizontal band of abstract, glowing red energy waves against a black background. The waves are composed of many thin, bright red lines that curve and flow from left to right, creating a sense of motion and depth.

# *Frequency* of data

# *Stability* of the solution





# *Location* of the data

# Looking for *questions*, or just *answers*?





***Power*** users, or ***business*** users?

# Which data can I *trust*?

“In essence, a *report watermark* is the visible evidence of a data governance process that enables organizations to combat spreadmarts and renegade reporting environments without unduly *repressing self-service* activity.”

YOU CAN  
CHOOSE  
TO SEE  
DATA  
DIFFERENTLY.



RED  
PILL  
ANALYTICS

[RedPillAnalytics.com](http://RedPillAnalytics.com)  
@RedPilla