Immersive Data Visualization

using Spark

Rosstin Murphy
Imran Younus
Jon Alter
IBM

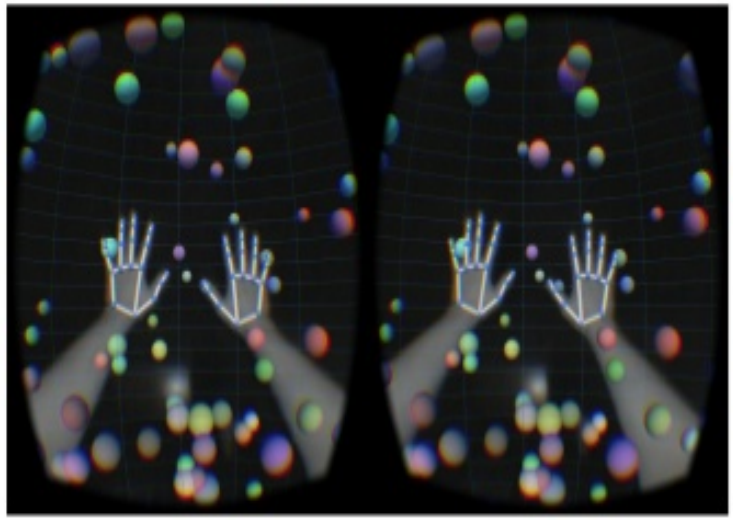
SPARK SUMMIT 2016
DATA SCIENCE AND ENGINEERING AT SCALE

Contents

- What is Immersive Data Visualization?
- Explanation of our Spark Data Pipeline
- Live Demo!
- Benefits of VR Data Visualization
- The Future VR Data Workbench

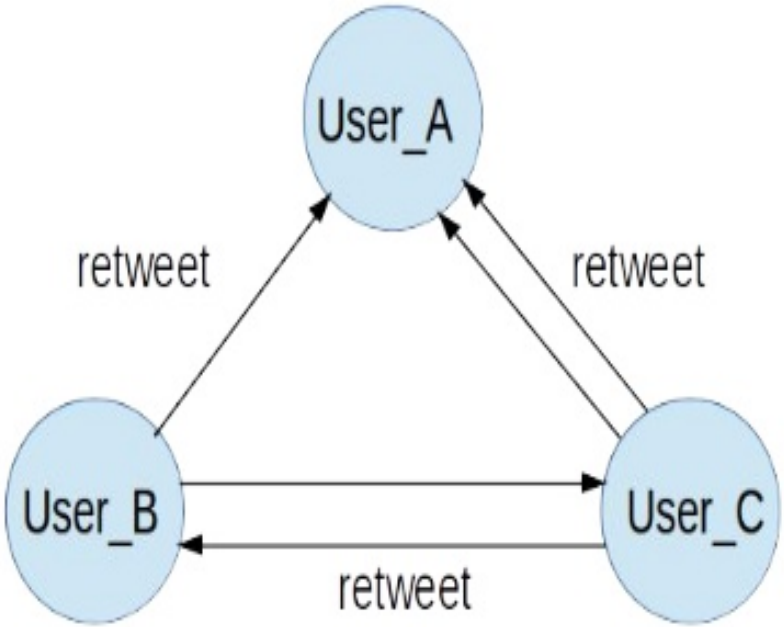
Immersive Data Visualization

- Uses a VR Headset such as the Oculus Rift to give you complete 360 degree vision
- Uses new 3D interaction methods, such as the Leap Motion infrared hand sensor.
- Data is represented in 3D and can be manipulated by the user's hands
- Users have depth perception, head movements, to fully perceive and manipulate 3D geometry



Retweet Graph from Single Topic of Discussion

- A conversation graph that represents activity related to a single topic of discussion
- We use Spark Word2Vec algorithm to obtain a set of related terms based on initial user query, which defines a topic



Retweet Graph from Single Topic of Discussion

- Collect all *retweets* containing top-*k* terms defining the topic
- This gives us a set of nodes and edges to generate the graph

	body	uid	ouid
61	Hillary says Bernie's plans aren't "realistic" or "pragmatic." This week she said "single payer ...	Original_Maven	MMFlint
969	Peeta is for @BernieSanders. Does that mean Bernie is team Peeta? Josh Hutcherson (@jhutch1992) ...	Chufi_Lalliter	cam_joseph
1459	@PokemonyeWest Would love to hear Bernie's thoughts on "Scott's Tots."	PokemonyeWest	john_hugar
1507	I'm in awe of @BernieSanders' energy & stamina. Age is only a number Boomers. Wow	jonathanstogner	mariashriver
1542	Thanks for trying to let @BernieSanders @SenSanders know how #SolarRoadways technology might fit...	woodymcfarley	SolarRoadways

edge list

Data Source: Twitter Decahose

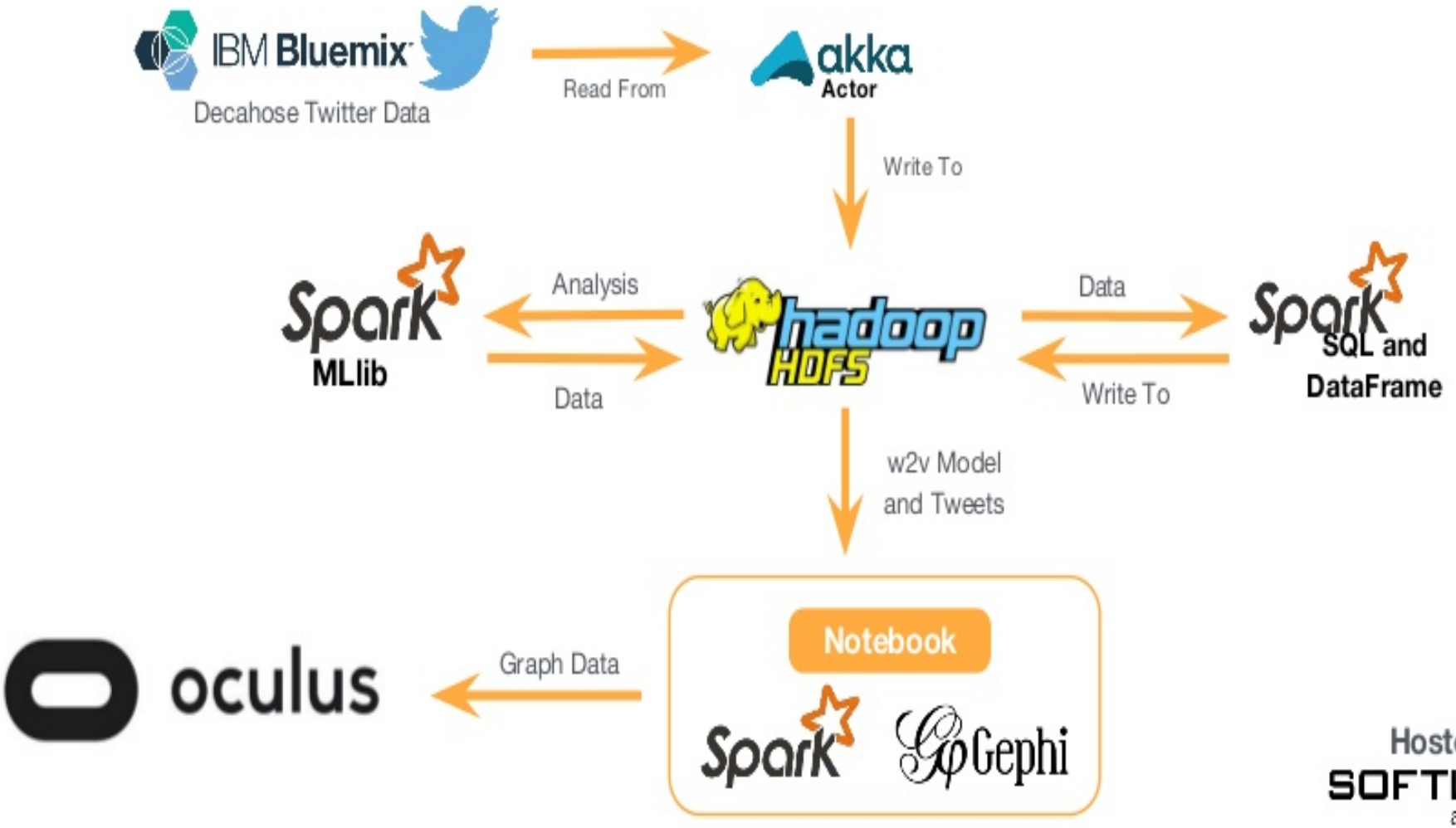
Retweet Graph from Single Topic of Discussion

- Community Detection Algorithms to detect communities of twitter user from retweet graph
- Force-directed algorithm for the layout
- Graph analytics algorithms to collect several attributes of the nodes in the graph

Edge List sorted by weight (# of retweets)

	uid	ouid	weight
0	padamo510	DrTomMartinPhD	9
1	GlennHeiser	InfamousGrace_	8
2	STERLINGMHOLMES	Libertea2012	8
3	lindamama02	Libertea2012	7
4	jkrupkin26	AidanKing_	6

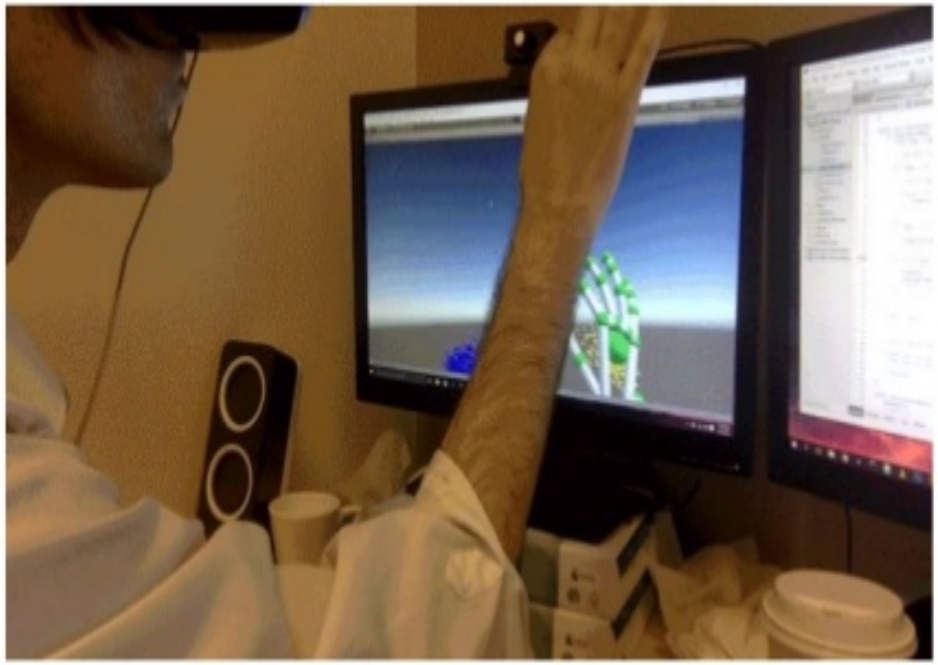
Application Architecture



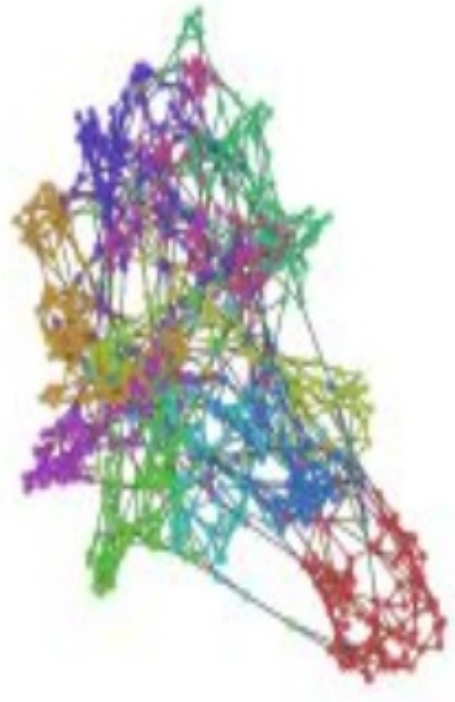
Hosted on
SOFTLAYER
an IBM Company

Benefits of VR Data Visualization - Demo

- Benefits of 3D
- Depth perception & perspective
- Gestures and convenience
- Infinite Screen Real-estate
- Sense of scale



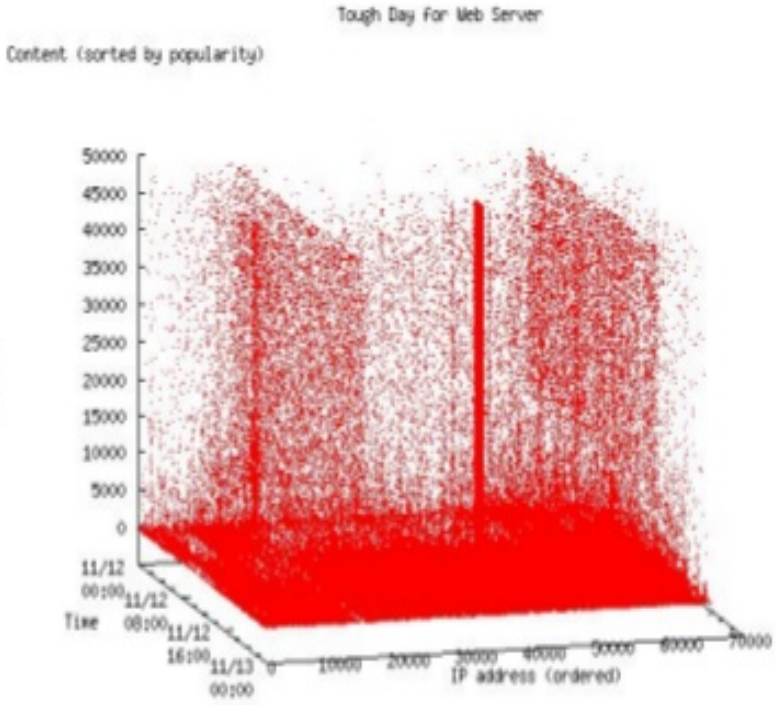
Benefits of 3D

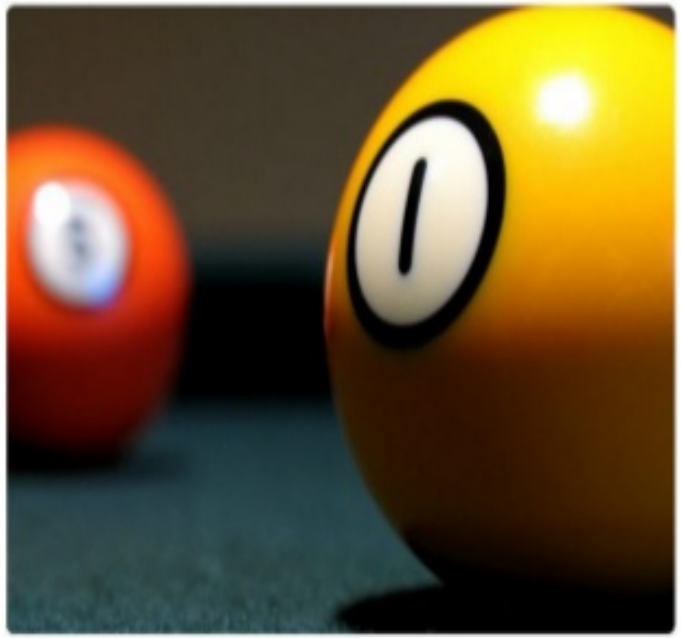


- Highly-dimensional data benefits from being 3D instead of 2D
- Some data, like geographic data, visualizes better in 3D
- for 2D graph: not possible to avoid line-crossing
- for 3D graph: can ALWAYS avoid line-crossing

Depth Perception

- You can represent a 3D image in 2D, but the depth relationships will not be clear
- With VR, micro-movements of your head reveal the relationships between elements
- 2 eyes = 2 views in VR, giving you the same depth perception you enjoy in the real world



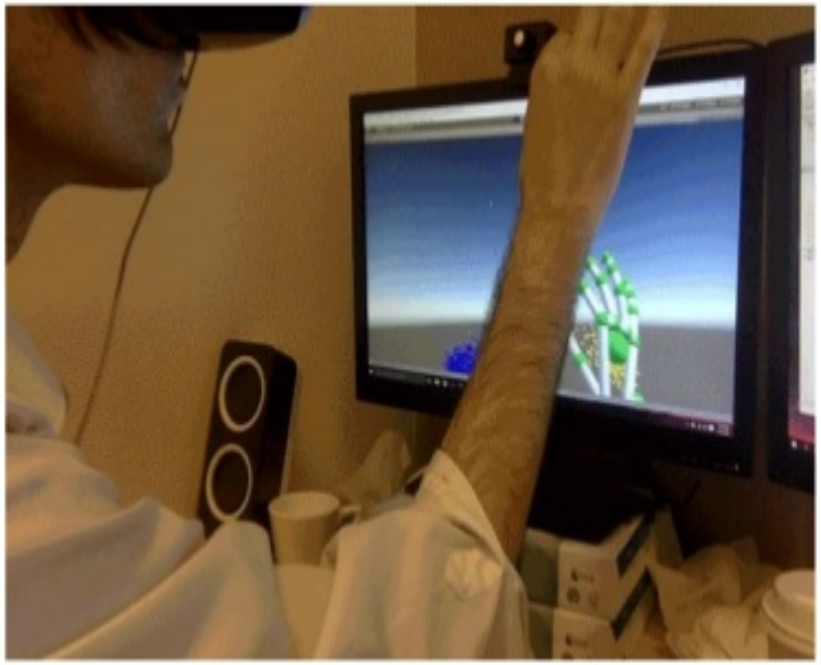


Infinite Screen Real Estate

- Many data scientists and traders invest heavily in multi-monitor setups
- The entire multi-monitor setup can be replaced with a single headset granting 360-degree vision

Gestures and Convenience

- Using hand gestures to manipulate 3D data is more natural than mouse
- As precision improves, this will become the preferred method of interacting with 3D objects
- You can still use keyboard and mouse for desktop interactions... but gestures can be a powerful tool for manipulating 3D visualizations



The Future VR Data Workbench

- Provide complete data analytics platform in VR, with abilities to perform exploratory data analysis and modeling along with immersive visualizations
- Multiple windows into the computer desktop controlled with keyboard/mouse
- Dynamically updated 3D visualizations of the data controlled with hand gestures
- Collaborative with multiple users

Thank you!

Rosstin Murphy - rosstin.murphy@gmail.com - @rosstinmurphy
Imran Younus - imranyounus@gmail.com
Jon Alter - jonalter@gmail.com