HOW TO BUILD A DATA STACK FROM SCRATCH

helpshift

Vinayak Hegde

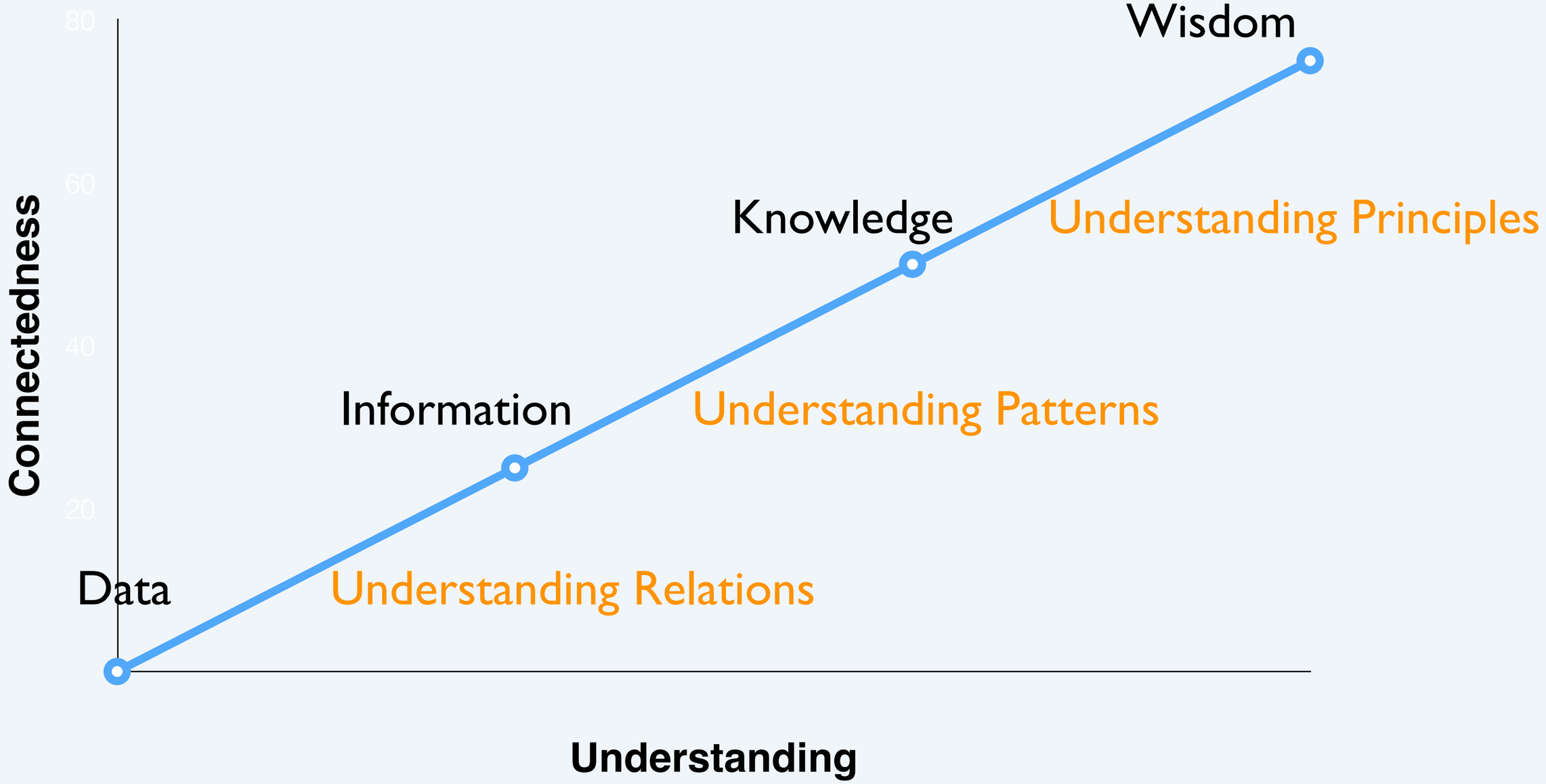
VP ENGINEERING

@vinayakh

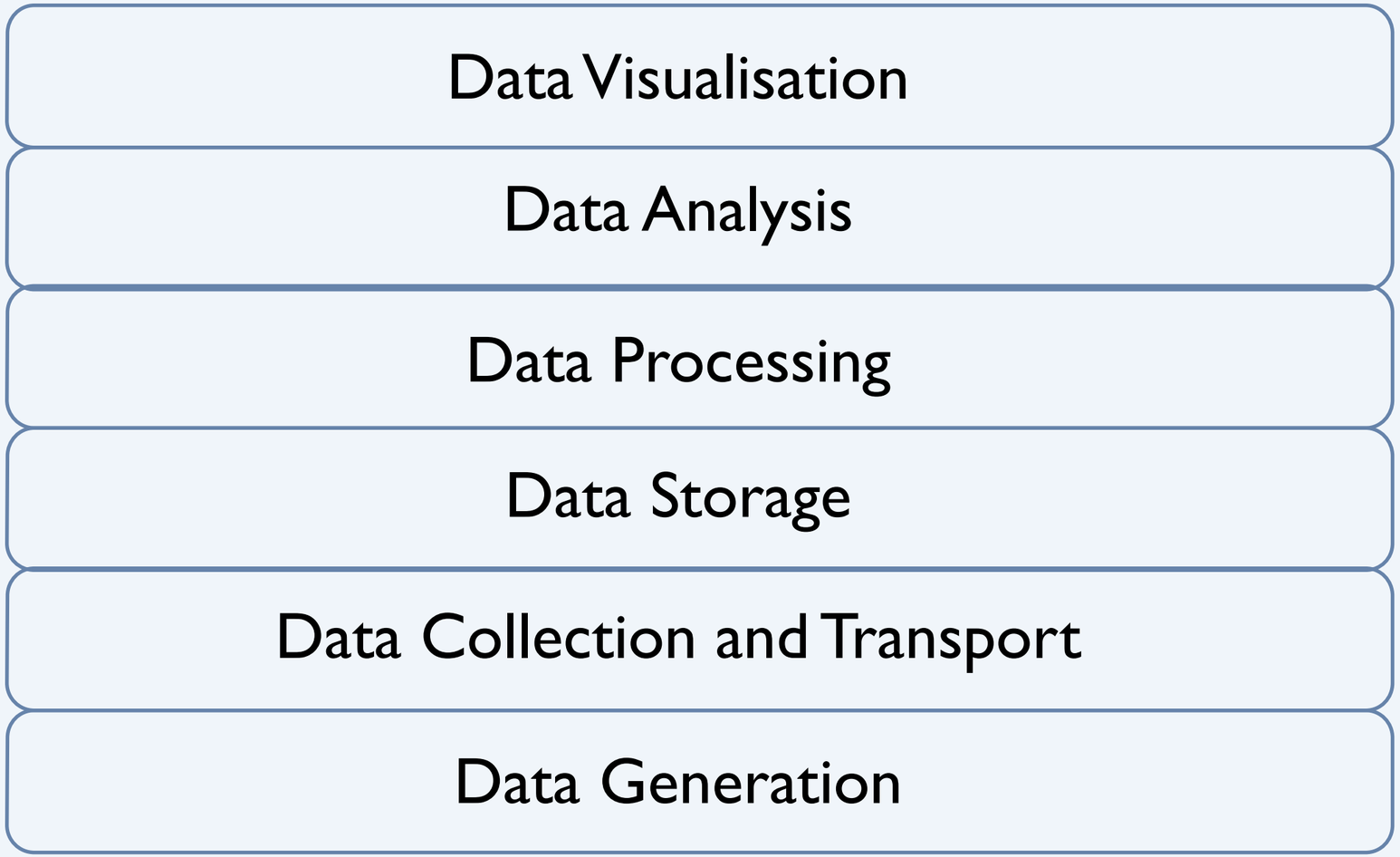
DATA IS THE NEW OIL

The Oil Metaphor

UNDERSTANDING DATA



THE DATA STACK

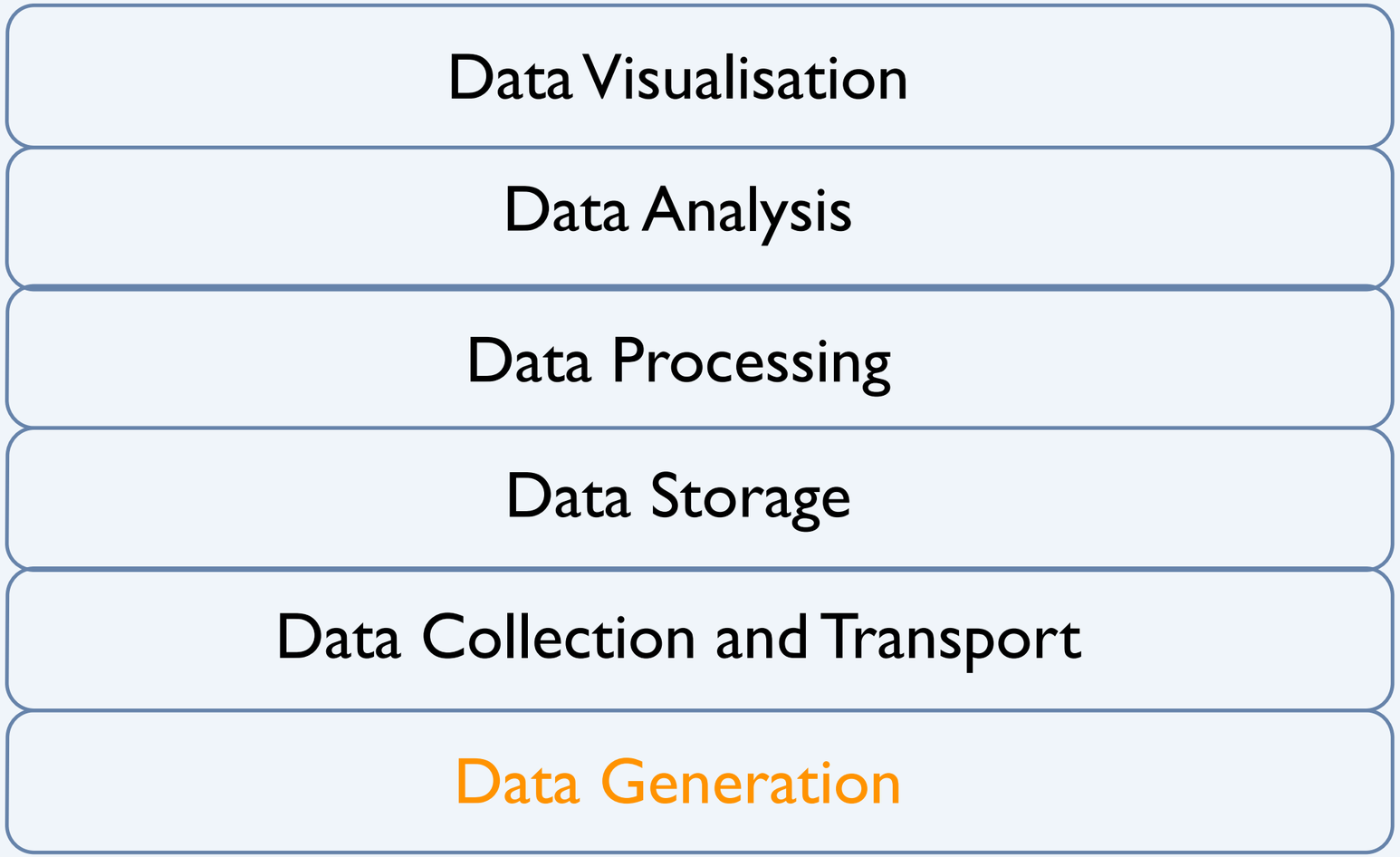


APPROACHES TO GETTING INSIGHTS

- Top-down Approach
  - Start with a hypothesis
  - Find data that can support or refute that hypothesis

- Bottom-up Approach
  - Look at nature of data
  - Look at the inter-relationships between different entities
  - Look at ratios, distribution, medians, variances,etc

DATA GENERATION



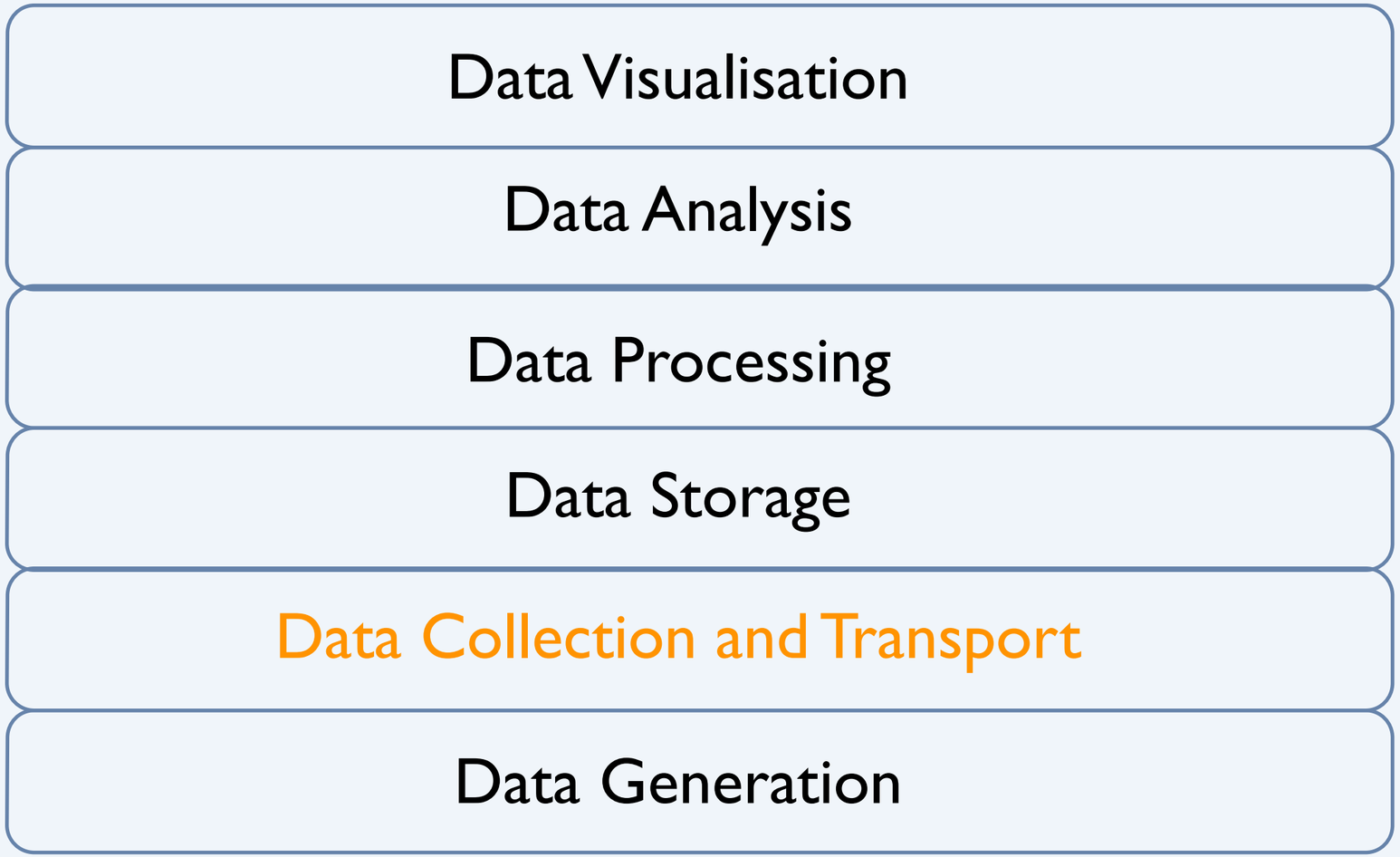
DATA GENERATION

- What data needs to be generated ?
- Frequency of generation
- Pre-aggregated or sampled
- Accuracy of data generation
- Is sample representative of population ?
- Format of data
- Metadata Enrichment
- Examples - Sensor reading, itemised store purchase data, Ad Impression data

DATA FORMATS

- CSV/TSV
- JSON
- Thrift file format
- RCFile

DATA COLLECTION AND TRANSPORT



DATA COLLECTION AND TRANSPORT

- Do some aggregation at source or send every data point
- Store locally and forward later
- Push Vs Pull methodology. Pros & Cons
- Factors in choice of underlying transport protocol
- Factors in choice of software
  - Reliability
  - Delivery policy / semantics
  - Durability and Fault Tolerance

DATA PROTOCOLS

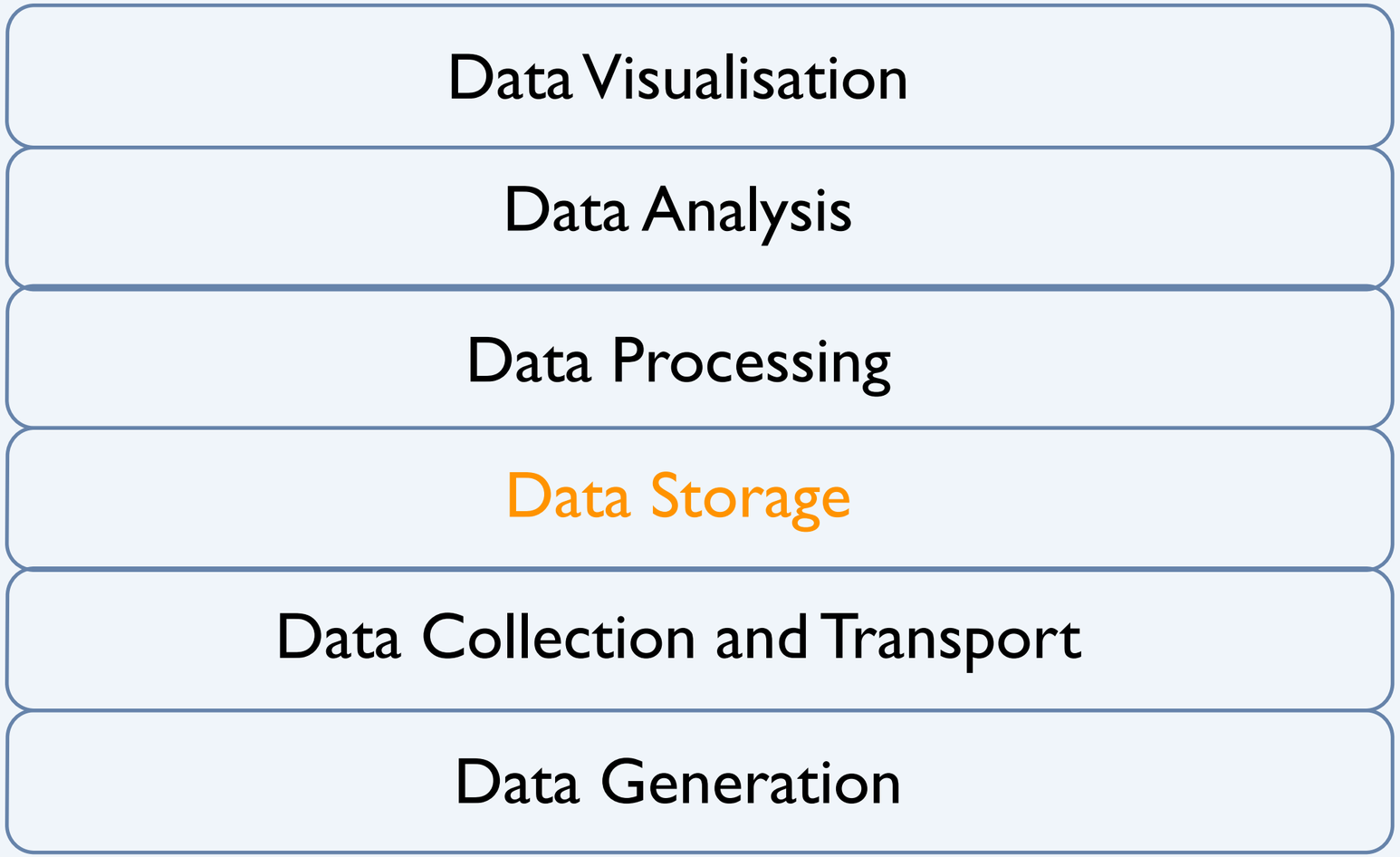
- TCP - connection oriented / reliable
- UDP - connection-less / unreliable
- MQTT - Useful for sensor data / resource constrained environments
- HTTP - REST APIs

QUEUEING AND ROUTING

Kafka VS RabbitMQ

Producer - Centric	Broker-Centric
Better for simple routing	Better if you want complex routing
Better for durable messages	Better for transient messages
More robust on failures of consumers/ At least once semantics	Many edge cases in which you can lose messages/ get replays
Better for larger message sizes	Better for smaller message sizes
More performant for large volume of messages	Performance can degrade with increase in message rates

DATA STORAGE



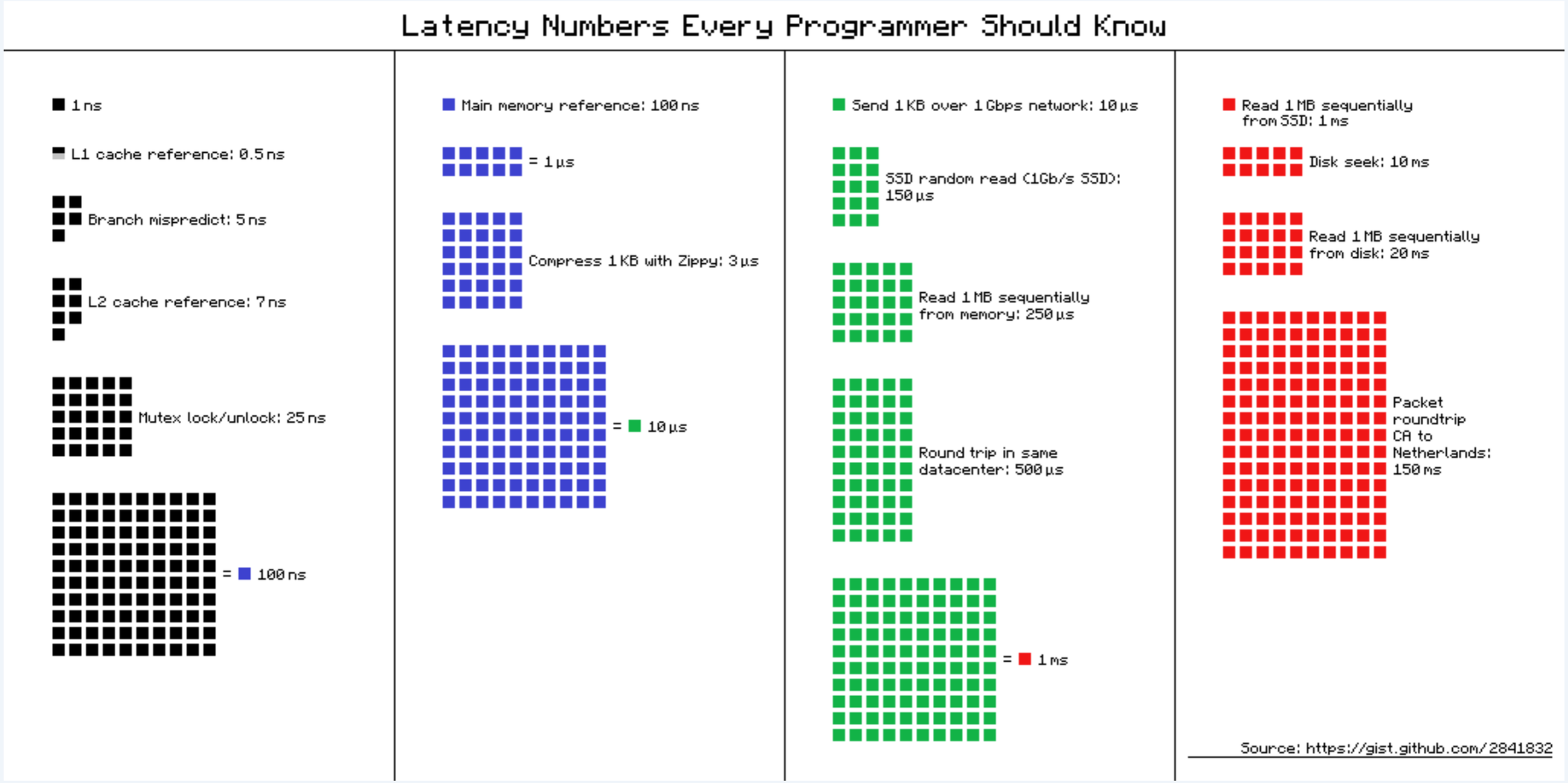
DATA STORAGE

- Storage media (SSD/Memory/Harddisk/ Network)
- Storage formats (B+Trees, Fractal Trees)
- Latencies of access
- Queryability and Indexes
- Filesystem differences

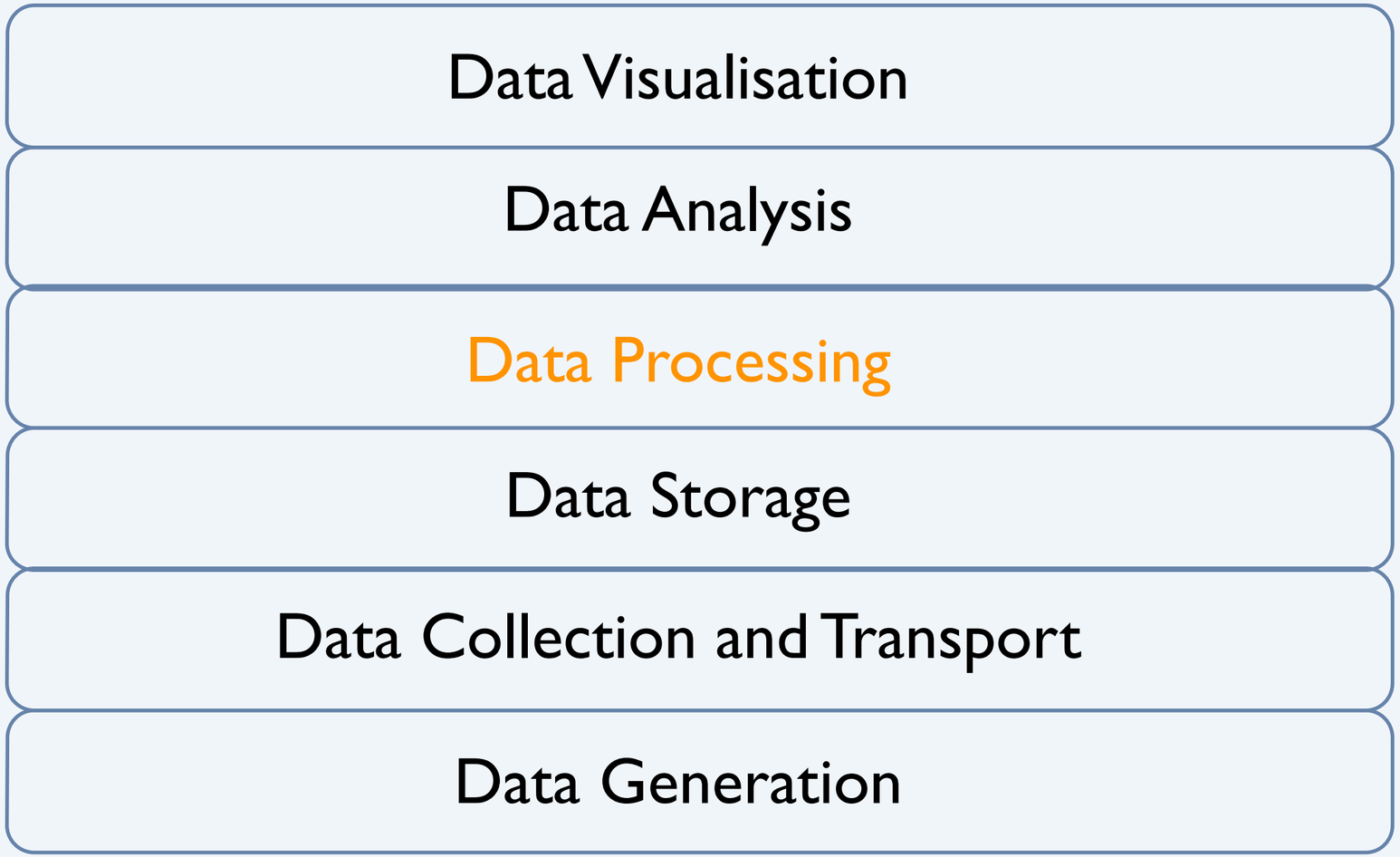
DATA ACCESS LATENCY

Operation	Time in ns	Comments
L1 cache reference	0.5	
Branch mispredict	5	
L2 cache reference	7	14x L1 cache
Mutex lock/unlock	25	
Main memory reference	100	20x L2 cache, 200x L1 cache
Compress 1K bytes with Zippy	3,000	
Send 1K bytes over 1 Gbps network	10,000	
Read 4K randomly from SSD	150,000	
Read 1 MB sequentially from memory	250,000	
Round trip within same datacenter	500,000	
Read 1 MB sequentially from SSD	1,000,000	4X memory
Disk seek	10,000,000	20x datacenter roundtrip
Read 1 MB sequentially from disk	20,000,000	50x memory, 20X SSD
Send packet CA->Netherlands->CA	150,000,000	

DATA ACCESS LATENCY



DATA PROCESSING



DATA PROCESSING

- Cronjobs
- Maps-reduce paradigms
- Message Processing Interface
- Iterative processing
- Microbatches
- Real-time Stream Processing

LAMBDA ARCHITECTURE

- Real-time stream processing
- Batch Processing

STREAM PROCESSING

STORM Vs SPARK Streaming

Task Parallel	Data-Parallel
Topology, Spouts and Bolts	RDDs
Good for working on individual items (record at a time)	Good for working on small groups of items (microbatches)
Sub-second latency	Few seconds latency
Good for data pipeline transformations (such as graphics or accumulators)	Good for iterative workloads (such as machine learning)
Resilent to failures of nodes (Nimbus,Zookeper)	Resilent to failures of nodes (Hadoop YARN, Mesos)
Fault tolerance - At least once	Fault Tolerance (Exactly once)

DATABASE - OLTP vs OLAP

MySQL Vs Infobright

Row based	Columnar Storage
Good for Transactional Workloads	Good for Analytical workloads
Low compression (size often increases by small factor on ingestion)	High Compression (size often decreases by huge factor on ingestion)
Loading is fast	Loading is slow and CPU intensive
Good for all kinds of data	Good especially for machine generated data
Sampling is possible but hard	Sampling and approximate queries are possible
Uses indexes and caching for better performance	Uses knowledge grid (metadata layer) for better performance

NOSQL DATA STORES

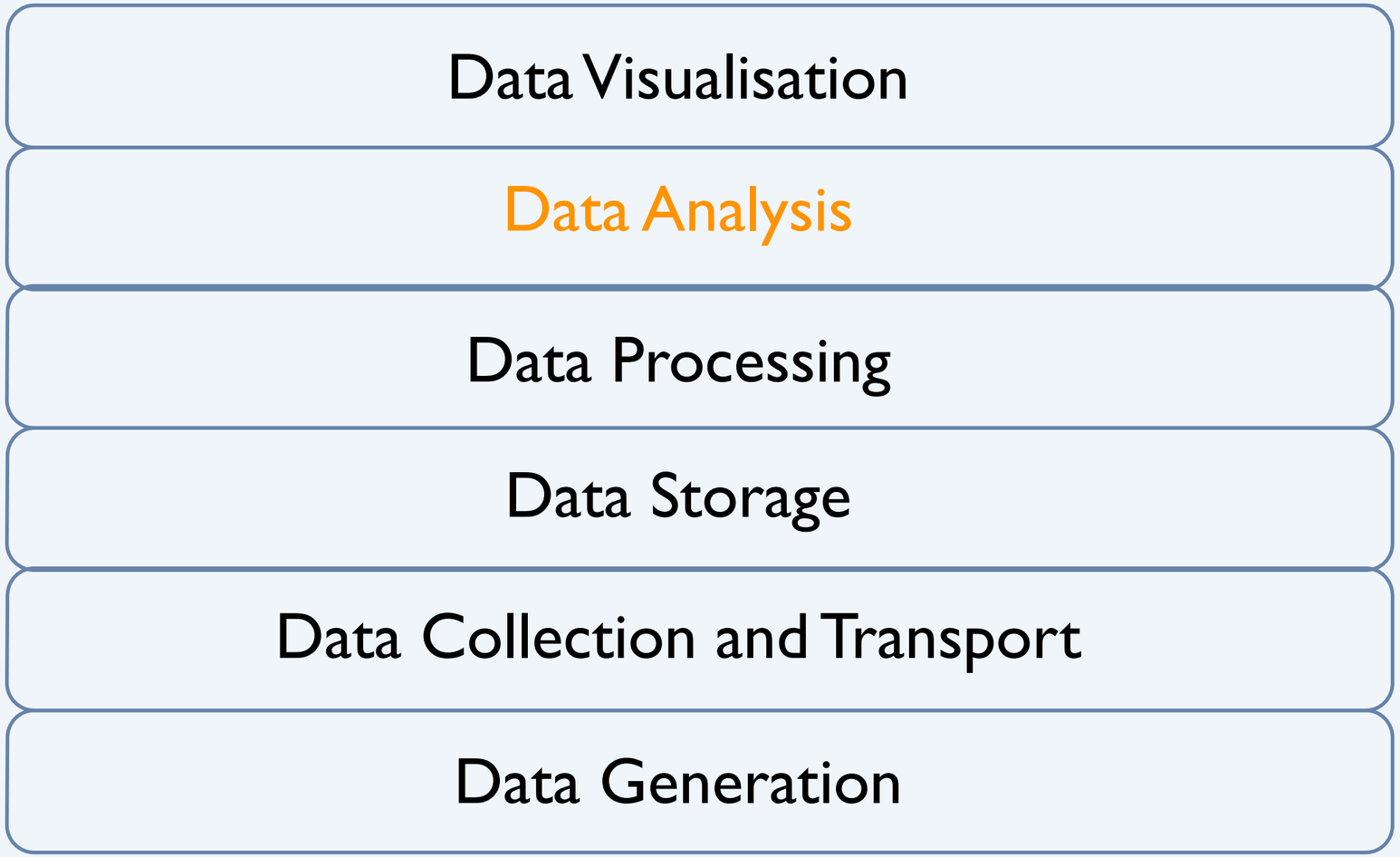
HBase Vs MongoDB

Wide-Column Store	Document Store
Schema-Free and no SQL support	Schema-Free and no SQL support
Has no types	Has types
Has no secondary Indexes	Has Secondary indexes
Has Triggers	Has no Triggers
Good Scalability due to HDFS	Decent Scalability but performance suffers
Selectable replication factor	Master-slave replication (though replica set can be large)

DATA LANDSCAPE

- Real-time processing systems (Storm)
- Complex Event processing (Esper)
- Big data batch (Hadoop)
- Big data iterative (Hadoop, Spark)
- Columnar Storage (Infobright, Vertica, RCFile)
- Memory-optimised systems (SAP Hana, Spark)
- Graph DB systems (neo4J, GraphX)

DATA ANALYSIS



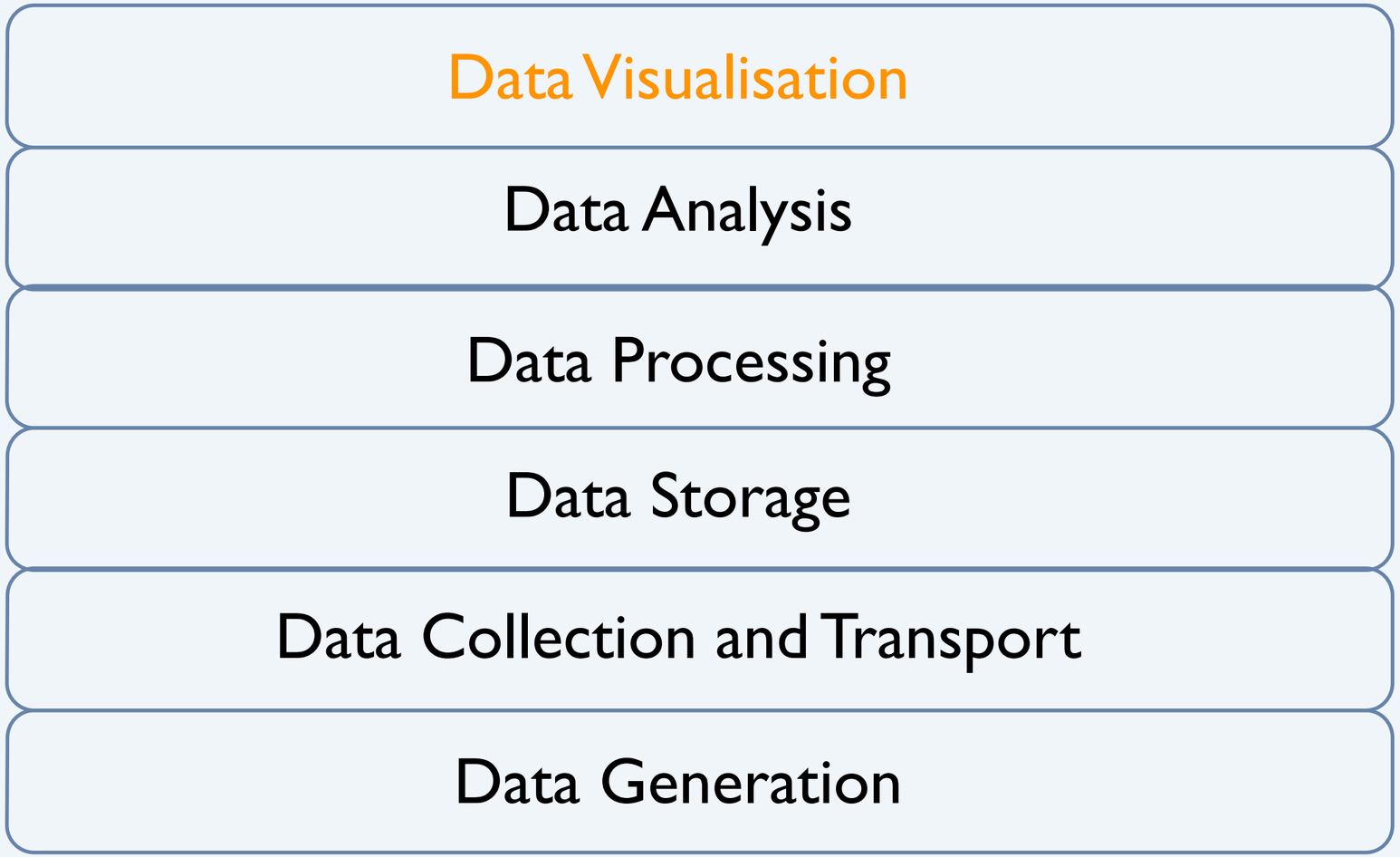
DATA ANALYSIS

- Merge metadata
- Layer 3rd party data
- Geocoding
- Aggregation
- Incorporate human input

DATA ANALYSIS

- Statistical analysis
  - Basic - Mean, Median, Variance, distribution, Outliers, Quantiles
  - Predictive models, Latent variable models
- Machine learning
  - Supervised learning
  - Unsupervised learning

DATA VISUALISATION



DATA VISUALISATION

- Visual cognition
- Visualisation as a narrative
- Color Palette
- Compare and contrast
- Find outliers and do exploratory analysis

Examples

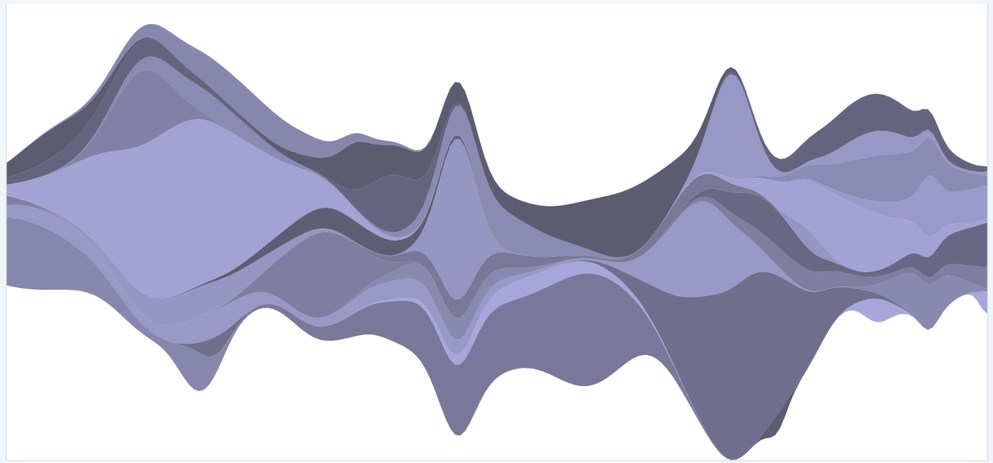
- Sunburst
- Stream Graphs

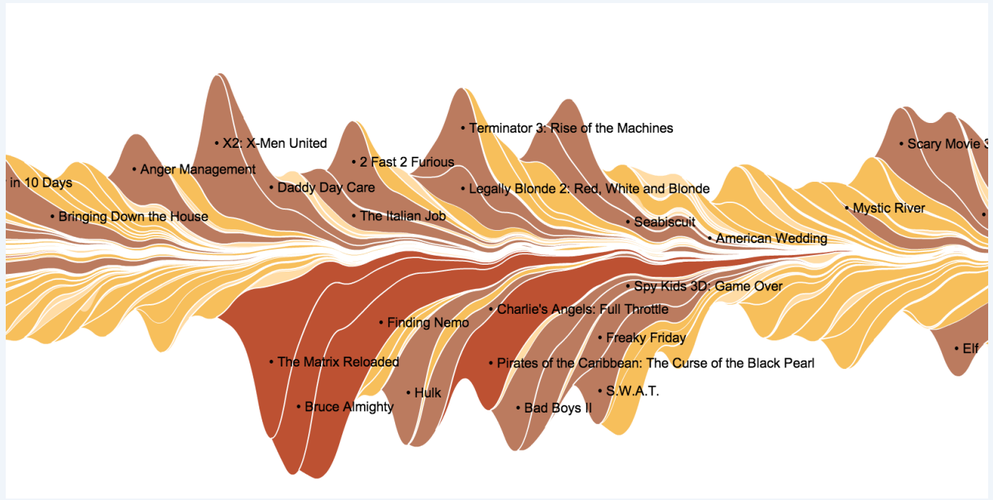
SUNBURST

Hierarchy + Relative Size

- Disk space usage
- Relative populations of Administrative blocks
- Market caps of sectors and companies listed on Stock market Index

STREAMGRAPHS





Time Trend + Relative magnitude

- Listening trends (last.fm)
- Topic streams (Twitter)
- Box office receipts of Popular movies (NYTimes)

LINKS

- https://thrift.apache.org/static/files/thrift-20070401.pdf
- http://json.org/
- http://mqtt.org/
- http://www.quora.com/RabbitMQ/RabbitMQ-vs-Kafka-which-one-for-durable-messaging-with-good-query-features
- http://www.akkadia.org/drepper/cpumemory.pdf
- https://gist.github.com/jboner/2841832
- http://stackoverflow.com/questions/24119897/apache-spark-vs-apache-storm

LINKS

- http://xinhstechblog.blogspot.in/2014/06/storm-vs-spark-streaming-side-by-side.html
- http://db-engines.com/en/
- http://bost.ocks.org/mike/algorithms/
- http://www.wired.com/2014/04/tree-diagrams-the-most-important-data-viz-tool-in-history/
- http://www.nytimes.com/interactive/2008/02/23/movies/20080223_REVENUE_GRAPHIC.html

QUESTIONS

- Blog
https://engineering.helpshift.com

- Twitter
@vinayakh

- Email
vinayak@helpshift.com

- Jobs
jobs@helpshift.com