

# Turning Any Data Into Consumable Analytics in the Cloud

Ian Thomas, Talend

# Analytics is Core to All Digital Transformation Initiatives

## Operational Effectiveness

How efficient are my leads?

Which channels are performing well?

How is web traffic trending when it's sunny?



## Customer Centricity

Are my customers responsive?

Is our brand popular?

How are we perceived amongst our competition?

Embedding intelligence across your business

# The New Data Challenges to Becoming Data Driven Organizations

## Ingestion

Various sources, multiple systems, diverse format, immediate needs

## Integration

Complex pipelines, data structures, specialized skills and resources

## Analysis

Slow, incomplete and limited access to data

## Collaboration

Conflict and a lack of collaboration between IT & Business, limited data sharing



# Data Lakes... A Path To Becoming Data-Driven

## But Data Lakes Create New Challenges...

Not trustworthy, not governed, lack ubiquity

- Store more for less  
90% Of deployed data lakes will be useless, Gartner
- Centralizing information
- Drive deeper insight  
32% of US firms is inaccurate, Experian

“Data lakes fail when they lack governance, self-disciplined users and a rational data flow,” Gartner

# Build a True Cloud Data Warehouse



Unite Diverse  
Data Sources

Safeguard  
through Governance

Achieve the  
Highest Data Quality

Enable  
Self-Service



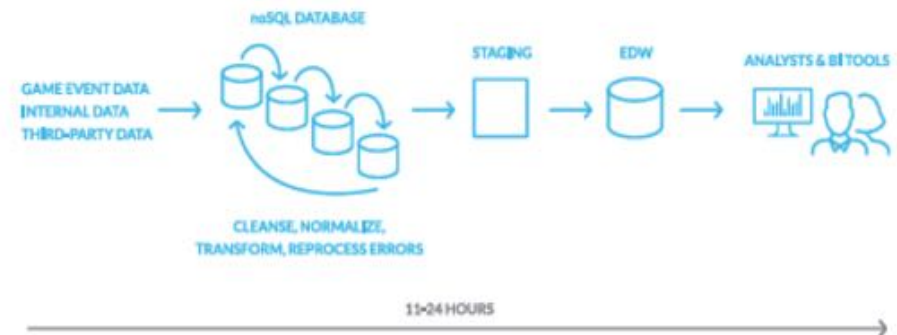
# Talend and Snowflake - Big Data Use Case

## NoSQL/Hadoop: Before

### Challenge

Take continuous data feeds from their games (3.5TB / day) and integrate that with other data into a holistic representation of game activity, usability and trends.

- Game event data stored in JSON format
- NoSQL DB used for processing – did not scale and data loss
- Long lead time to visibility
- Costly to maintain NoSQL DB



# Talend and Snowflake - Big Data Use Case

Snowflake: **After**

## Solution:

Take in the data from Amazon Kinesis, load it into an Amazon S3 landing area. Talend process runs every 5 minutes and then loads the files directly into an event log table in Snowflake, which makes all the JSON attributes queryable.

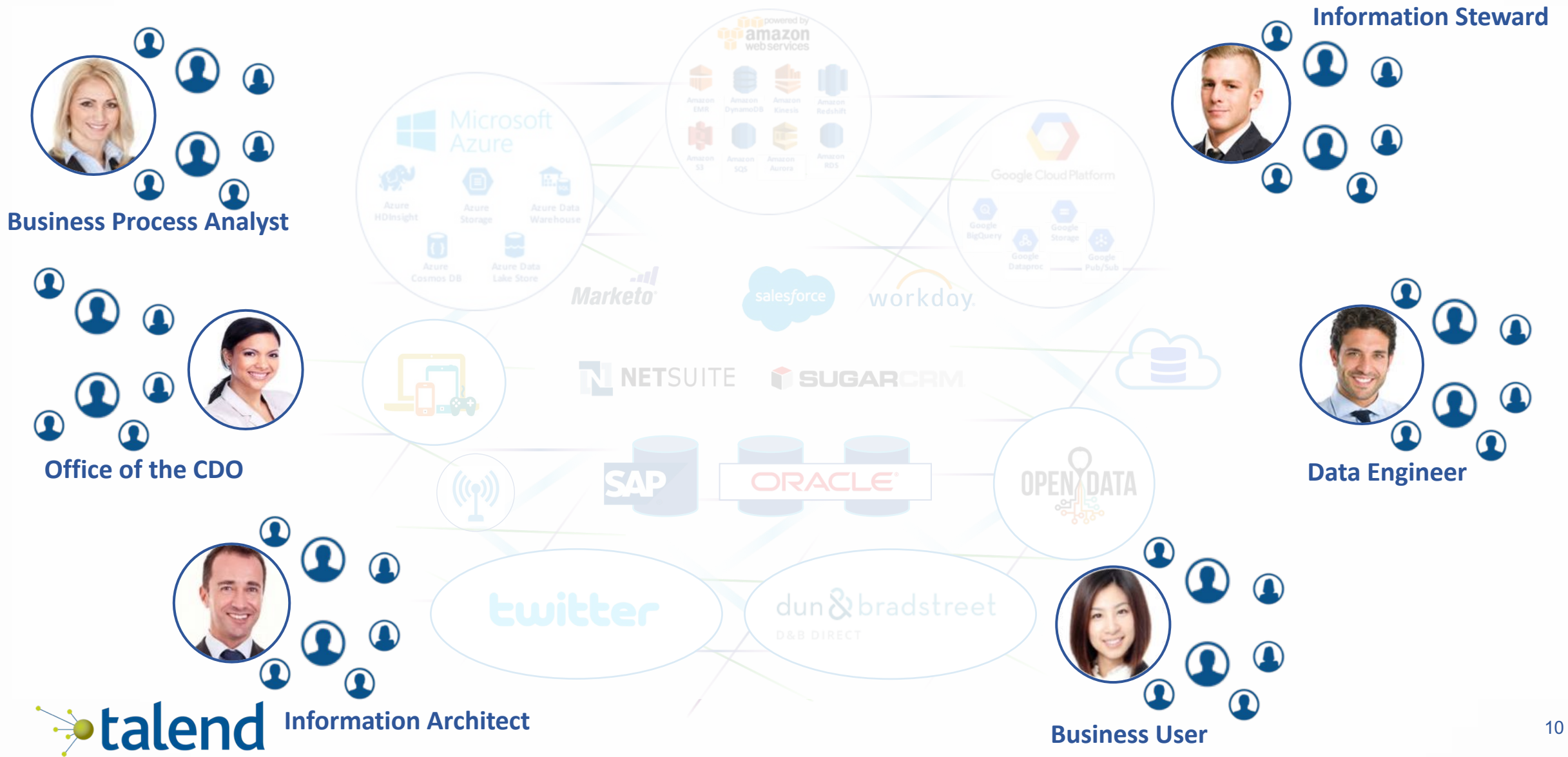
- 132x improvement in time to visibility (~5 min)
- 100% reduction in data loss
- SQL based access to historical data
- Lower TCO
- 100% cloud within next year



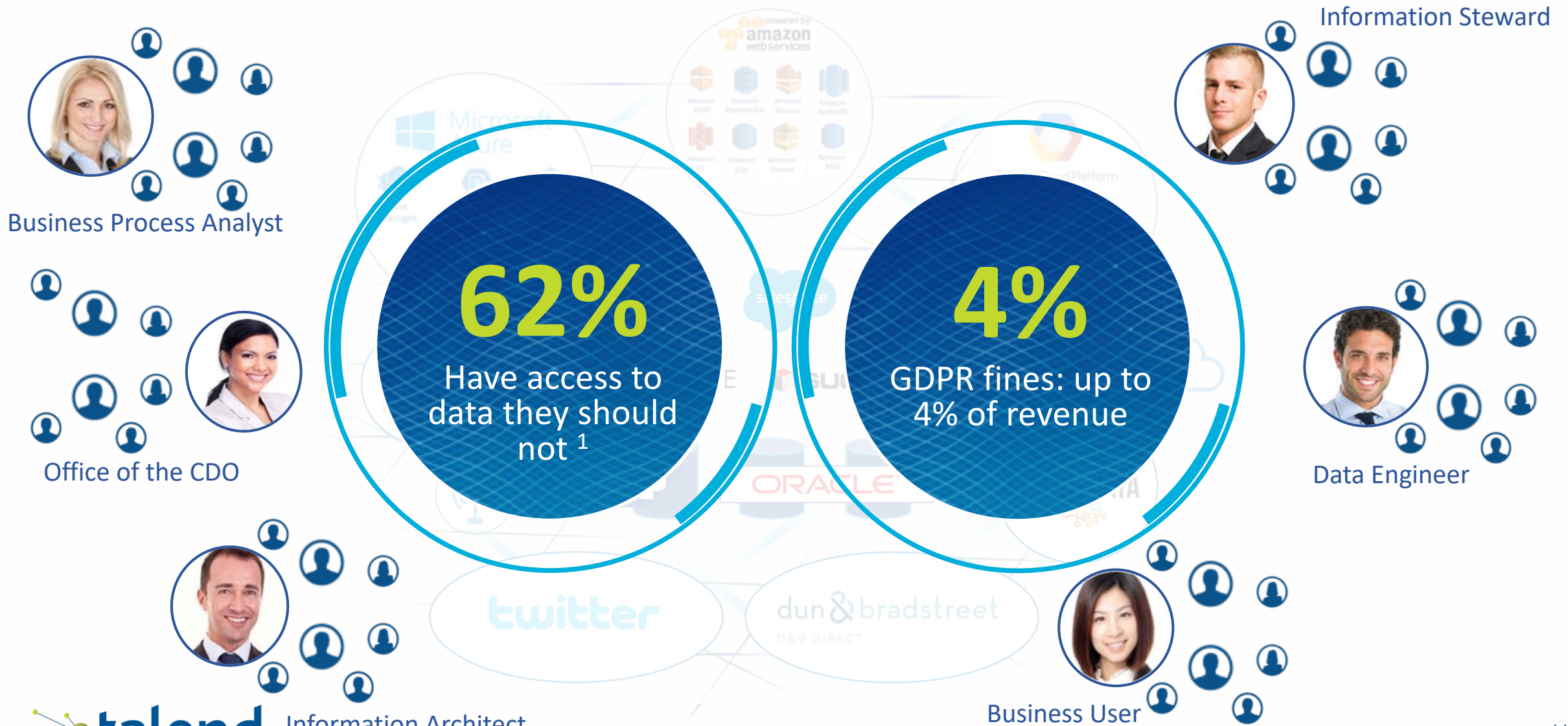
# An Explosion of Data ...



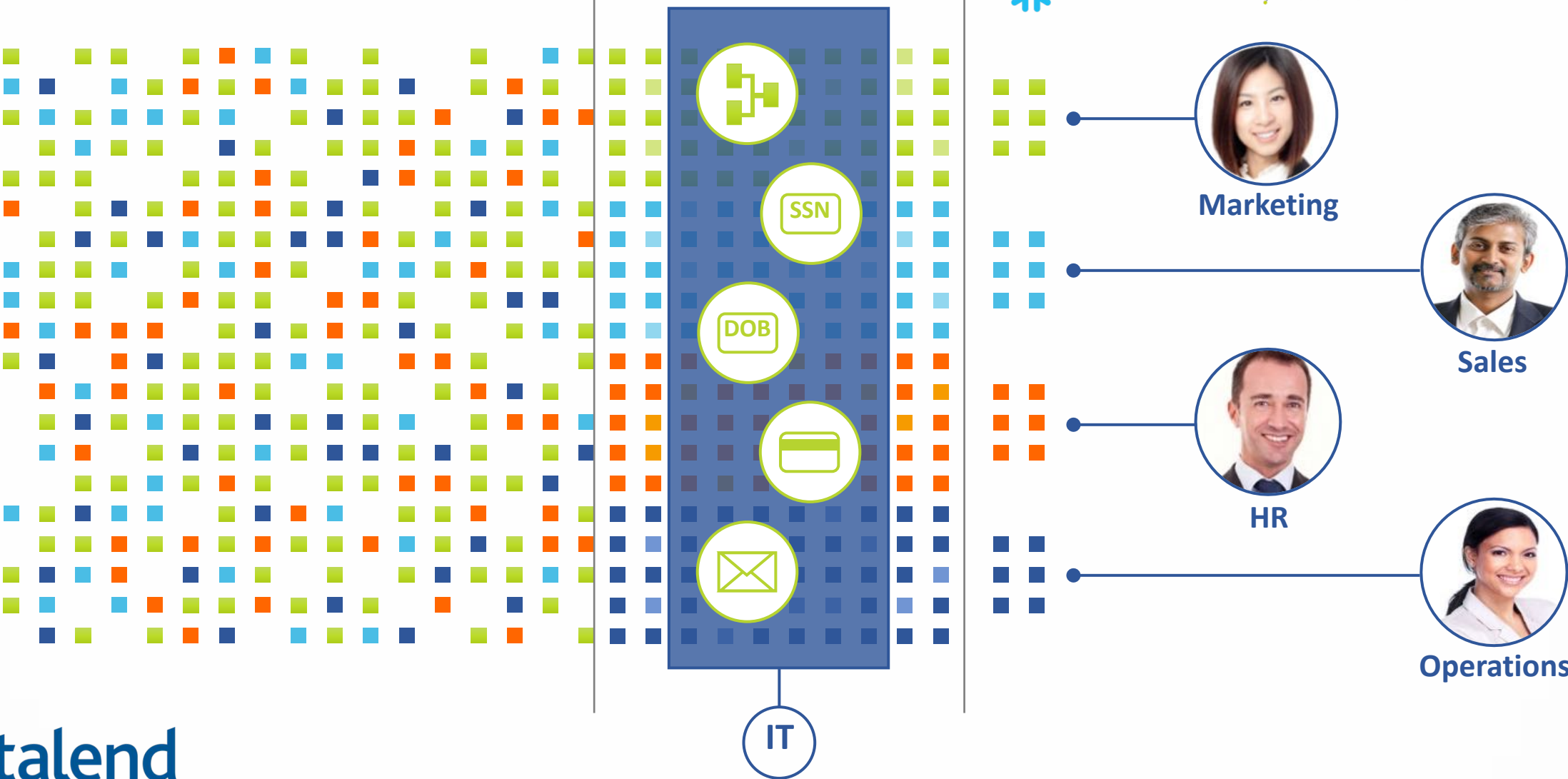
# ... More People Need Access To Data ...



# ... Lack of Governance Puts Enterprise Data at Risk ...



# Self Service – Access Data From Any Where





A premier provider of healthcare analytics, data and insight products to the world's leading pharma, biotech and medical tech companies.



## An IT Architecture Use Case

# DRG Big Data Initiative

## Challenges

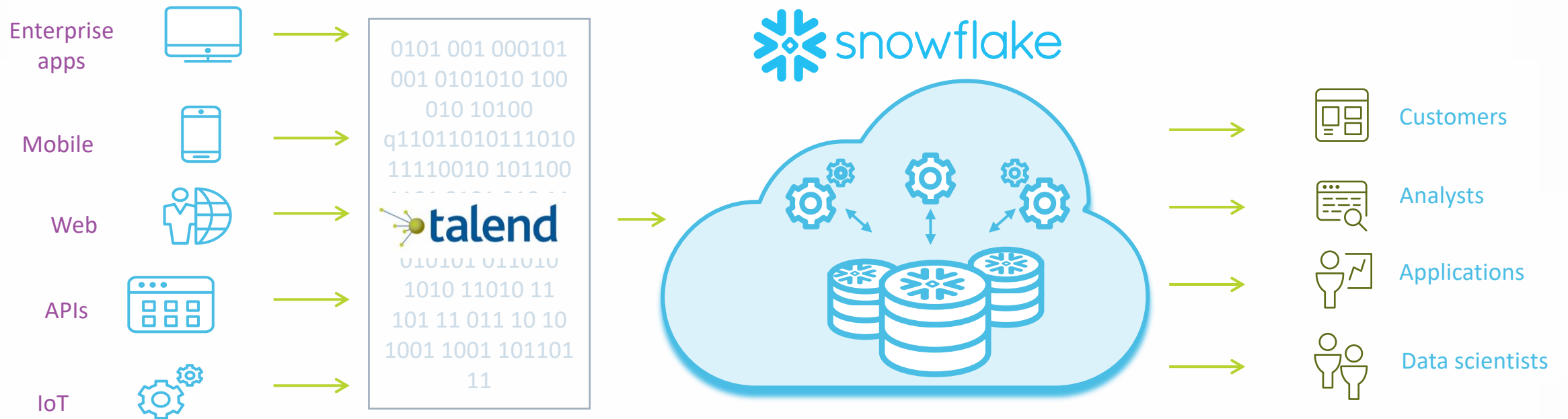
- Infrastructure was not ready for big data
- Building a big data team was costly and takes long
- Existing big data solutions were complicated & hard to integrate
- Pressure to move fast

## Solution

- A mature SQL Engine that works with big data
- Supports multi-terabyte data volumes
- Hosted in the Cloud
- Robust connectivity and Access



# Faster Cloud Analytics for Diverse Data



- 20X performance boost for bulk loading data
- Parallel data processing
- Easily move on-premises and cloud data to Snowflake