

# Big data e sicurezza nelle banche: l'estensione delle regole di governance a tutte le informazioni sensibili



**Sergio Mucciarelli**  
Data Security Leader – IBM Italia

---

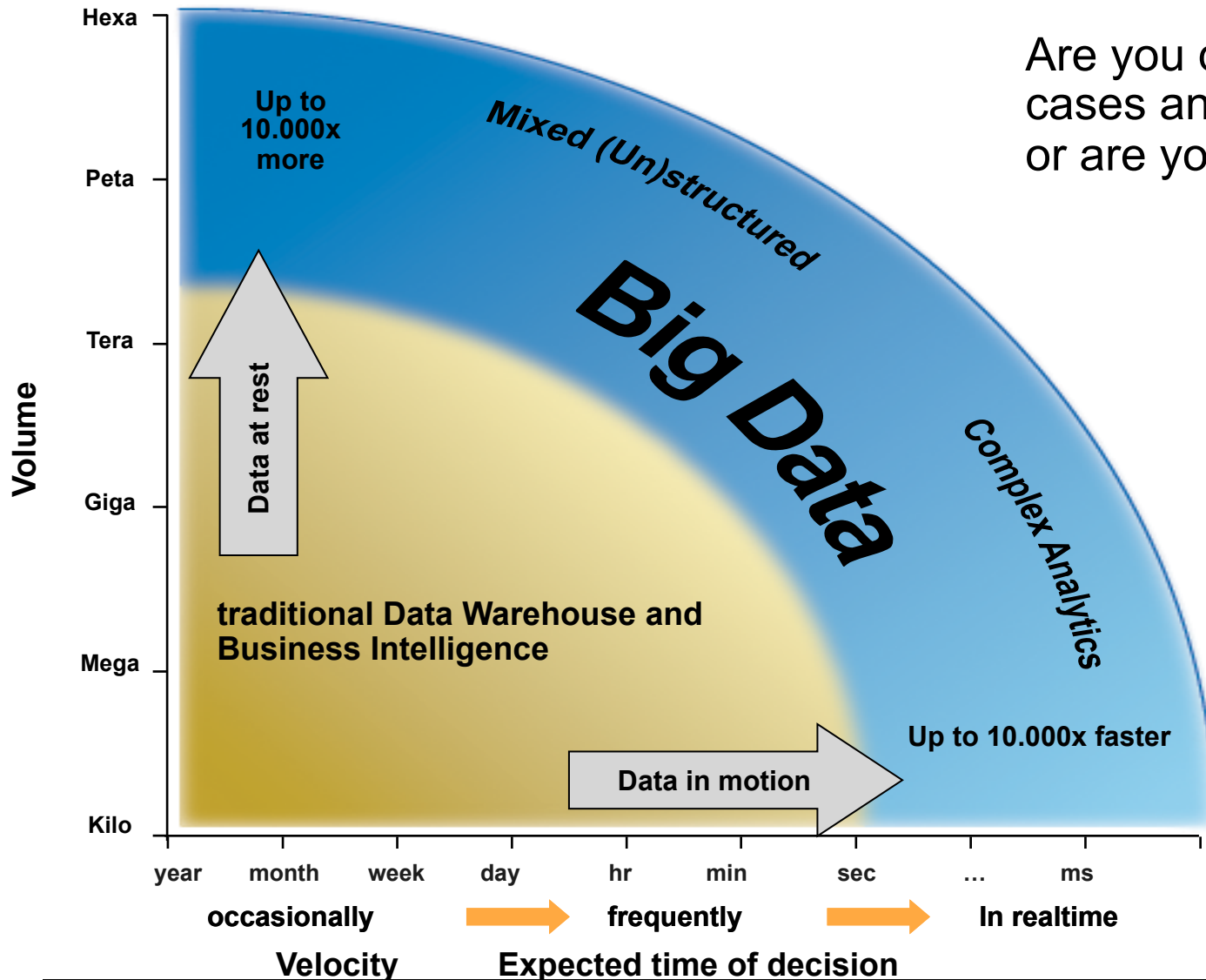
## Agenda

- **Big Data: new capabilities for banking**
  - **Extending Data Governance Security**
  - **An enterprise approach to monitor data activity**
  - **Big Data to enrich security intelligence**
-

# Big Data: new capabilities for banking



# Big Data offers new strategic business cases



Are you driving new use cases and new capabilities or are you driven ?

Up to 10,000x faster

occasionally

Data in motion

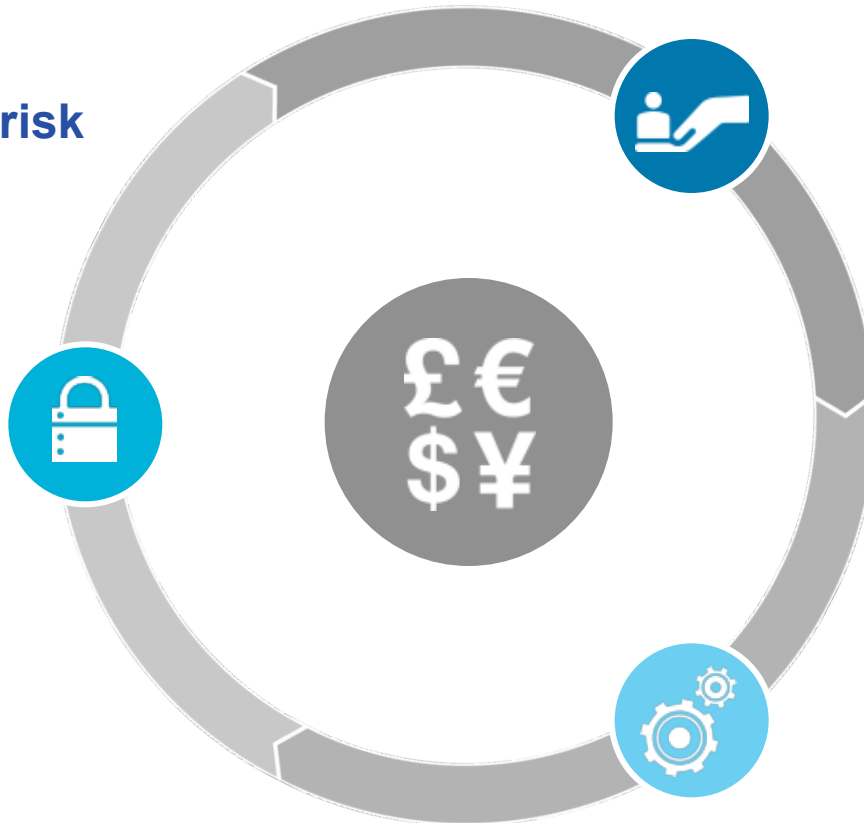
frequently

In realtime

Velocity

Expected time of decision

# Banking key big data business use cases



## Optimize enterprise risk management

- **Counterparty Risk Management**
- Real-time Fraud Detection
- AML & Fraud Investigation
- Financial Risk Analysis
- Market & Portfolio Risk Analysis
- Lending Risk Analysis
- Market Surveillance

## Create a customer-focused enterprise

- **Cross Selling & Offer Optimization**
- Contact Center Analytics
- Customer Retention
- Enhanced Customer Segmentation
- Data Monetization & Merchant Analytics
- Market Trading Analysis

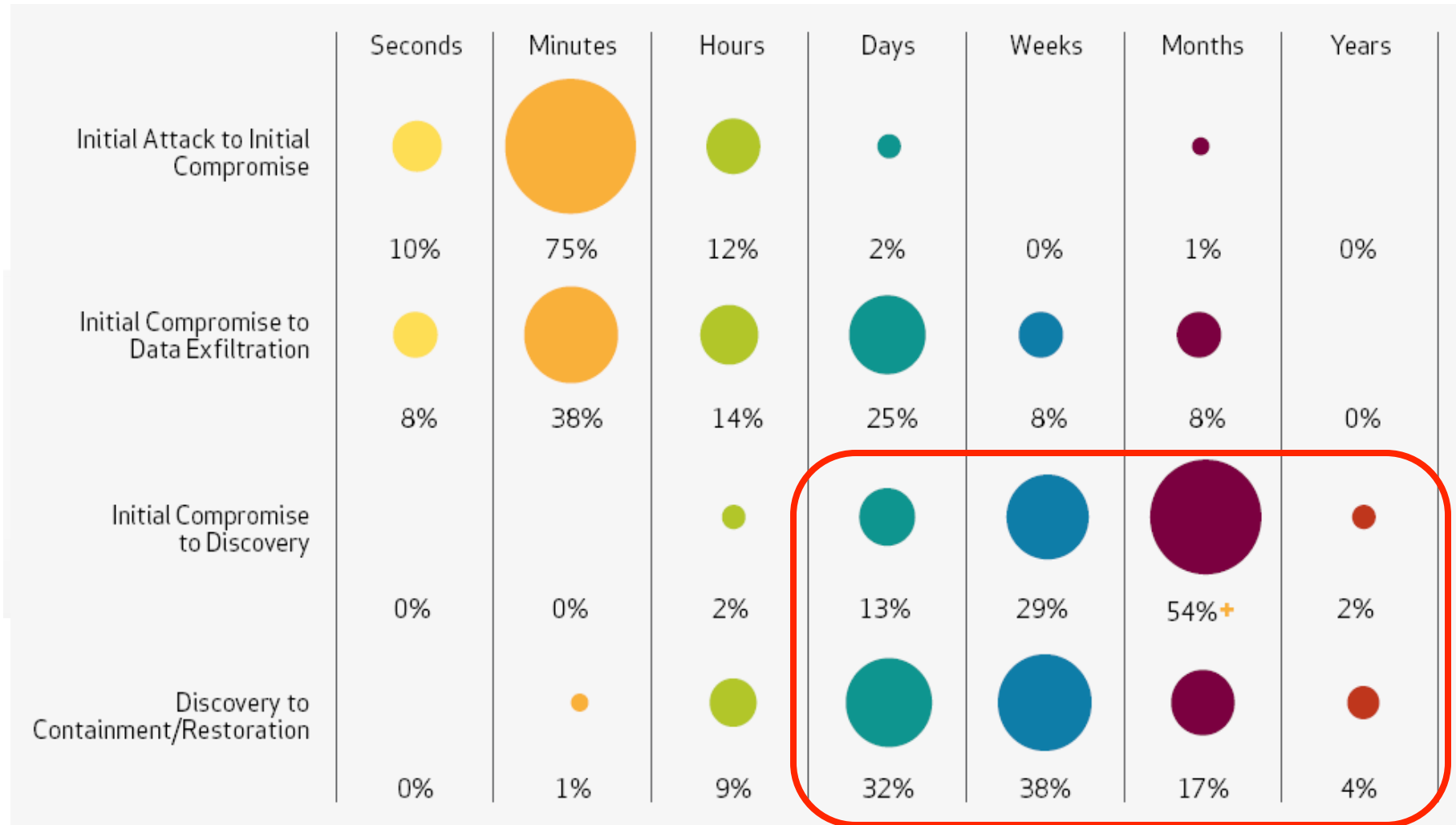
## Increase flexibility and streamline operations

- **DWH augmentation**
- **Security**
- Data Staging & Management
- System Log Analysis
- System Failure Analysis

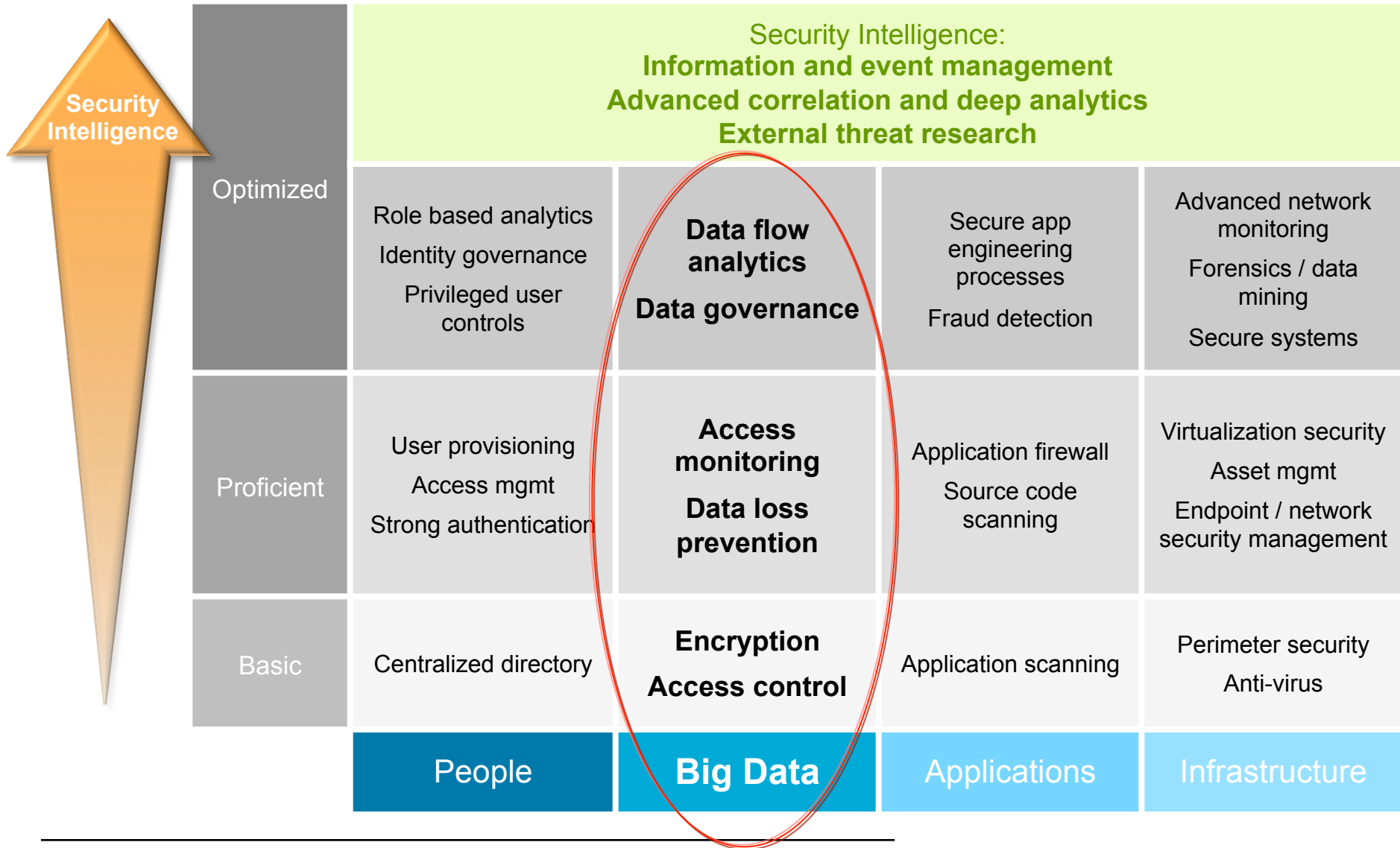
# Extending Data Governance Security



# Compromises take days or more to discover in 96% of cases; and weeks or more to contain in over 91% of cases



# Security Intelligence is enabling progress to optimized security





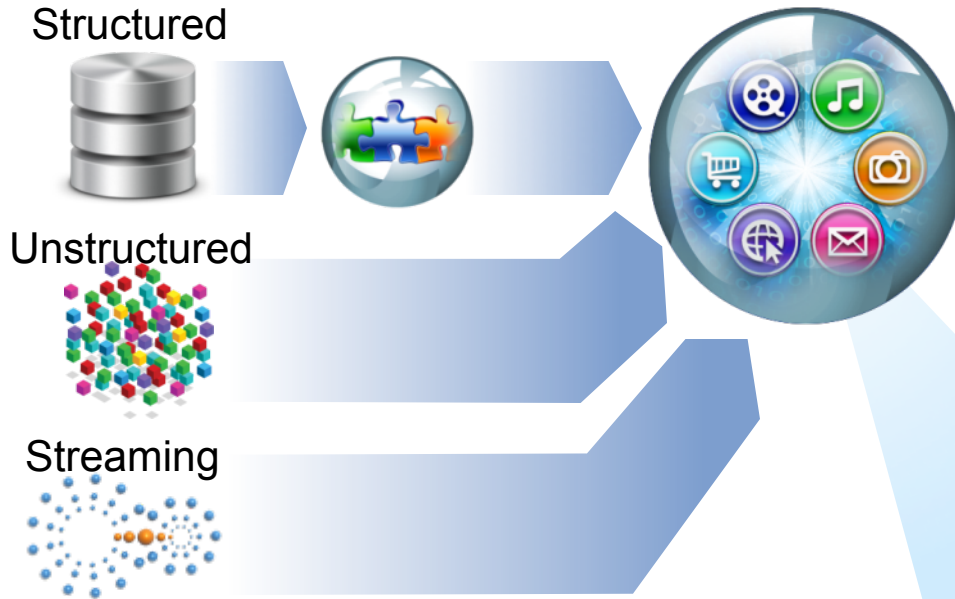
---

# An enterprise approach to monitor data activity



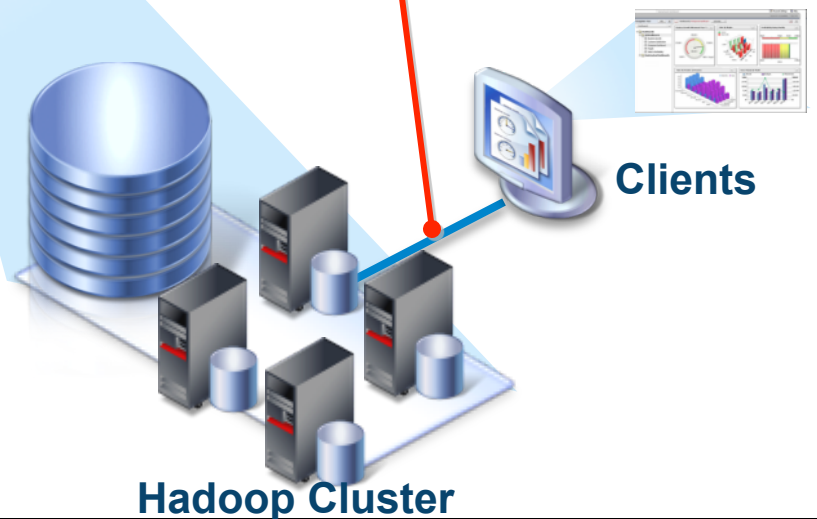
# Security and Compliance concerns are similar in Big Data

## Big Data Platform



- Who is running specific big data requests?
- What map-reduce jobs are they running?
- Are they trying to download all of the sensitive data for non-authorized purposes?,
- Is there an exceptional number of file permission exceptions?
- Are these jobs part of an authorized program list accessing the data?
- Has some new query application been developed that you were previously unaware existed?

- Massive volume of structured data movement
  - 2.38 TB / Hour load to data warehouse
  - High-volume load to Hadoop file system
- Ingest unstructured data into Hadoop file system
- Integrate streaming data sources

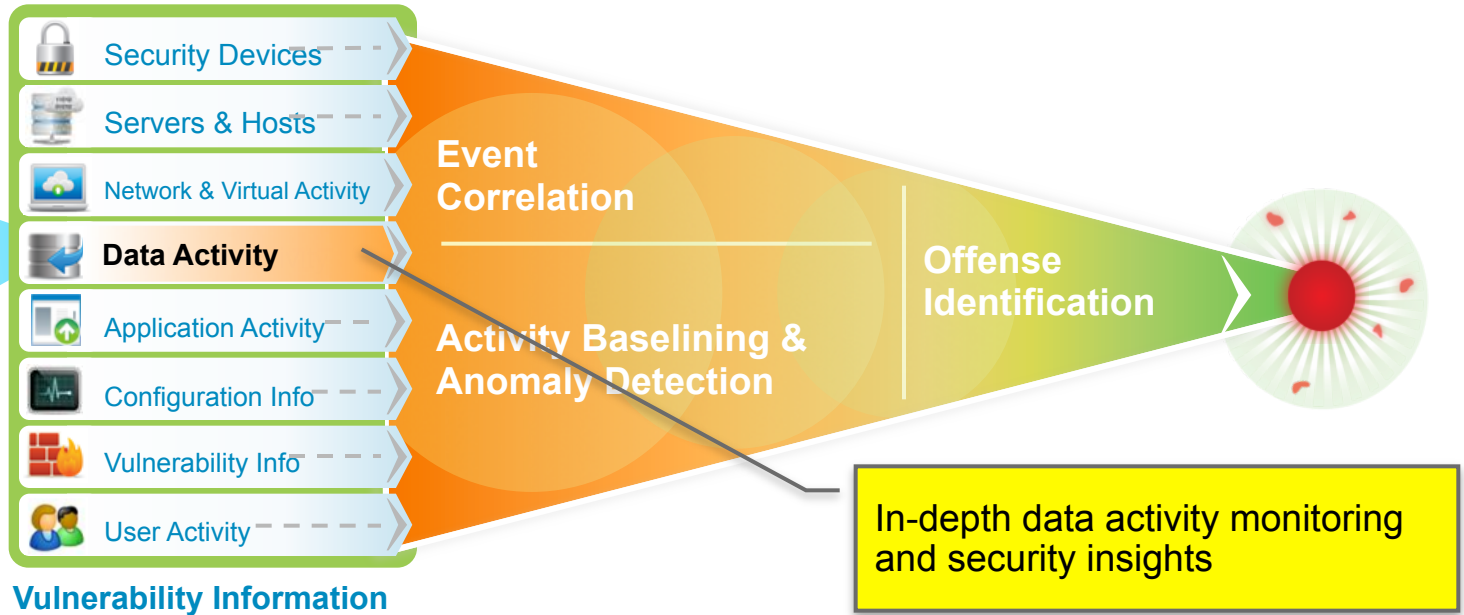


**Hadoop Cluster**

**Clients**

# Data security insights to your security intelligence

- Databases
- Data Warehouses
- Hadoop big data environments
- File shares



Extensive Data Sources + Deep Intelligence = Exceptionally Accurate and Actionable Insight

# Combat advanced threats with pre- and post-exploit intelligence and action



Vulnerability

PREDICTION / PREVENTION PHASE

Exploit

REACTION / REMEDIATION PHASE

Remediation



**Pre-Exploit**

**Post-Exploit**

## Prediction & Prevention

## Reaction & Remediation

Risk Management. Vulnerability Management.  
 Configuration and Patch Management.  
 X-Force Research and Threat Intelligence.  
 Compliance Management. Reporting and Scorecards.

Network and Host Intrusion Prevention.  
 Network Anomaly Detection. Packet Forensics.  
 Database Activity Monitoring. Data Leak Prevention.  
 SIEM. Log Management. Incident Response.



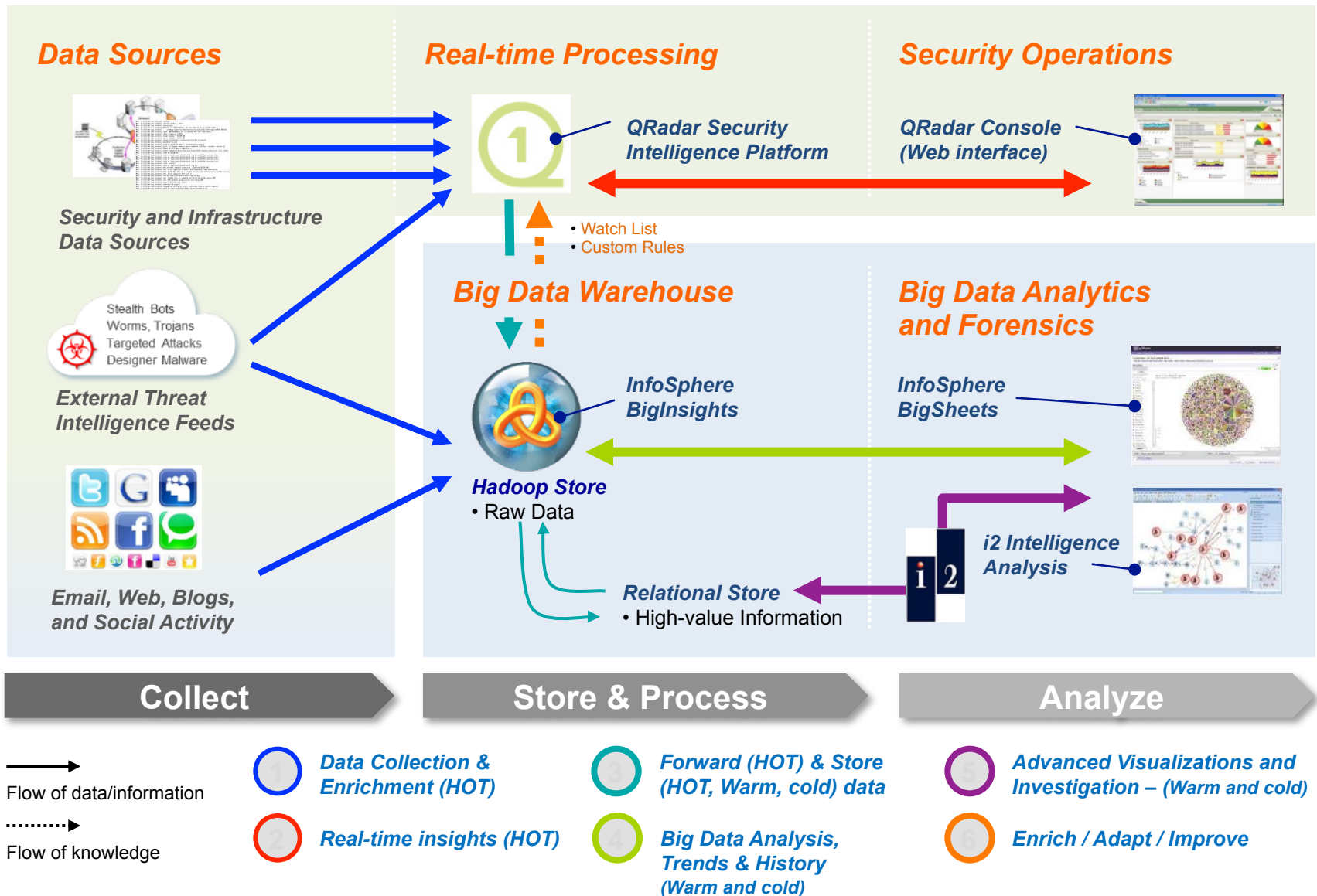
**IBM Security Intelligence**



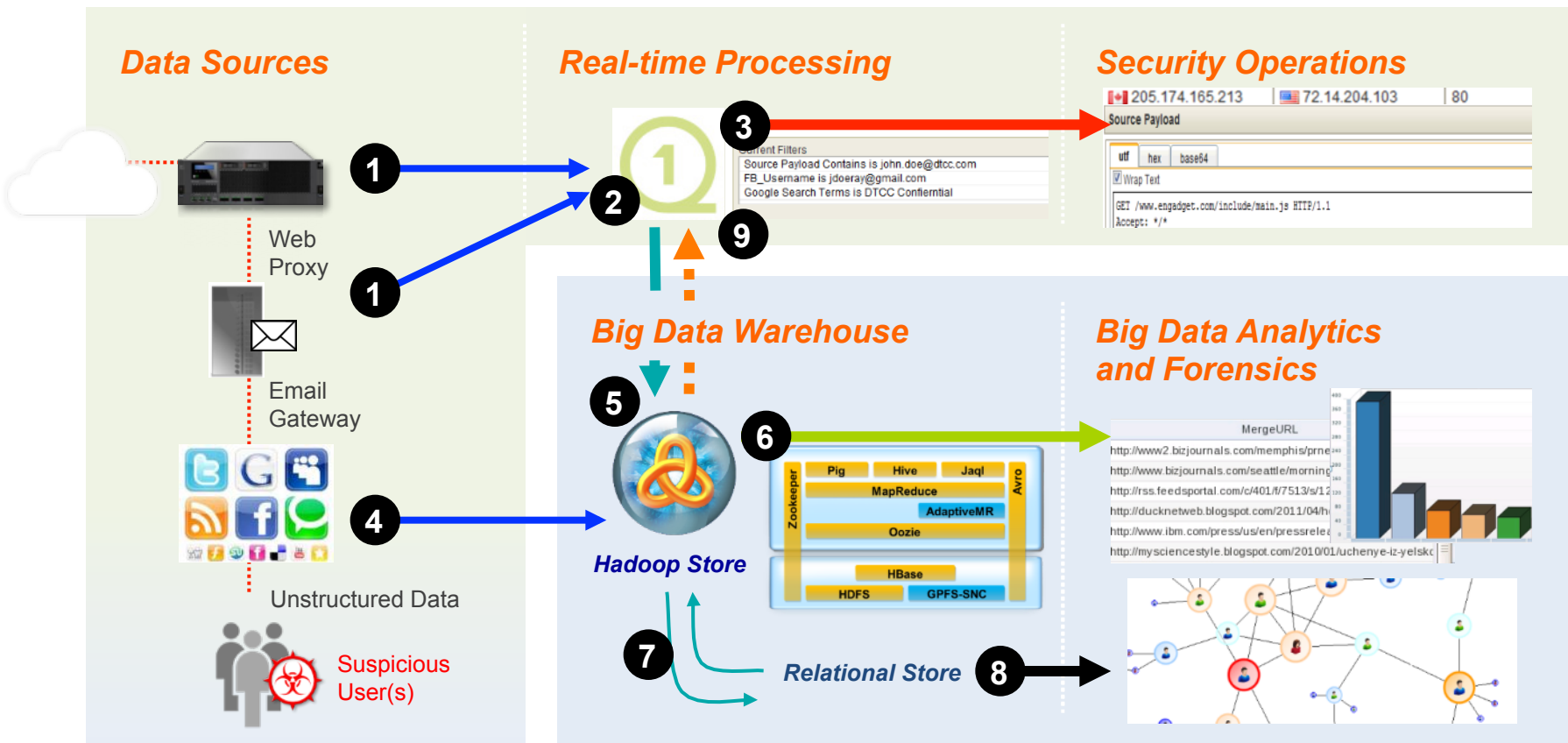
# Big Data to enrich security intelligence



# Security Intelligence with Big Data – Components and data flow



# Use case #1 – User profiling based on multiple sources



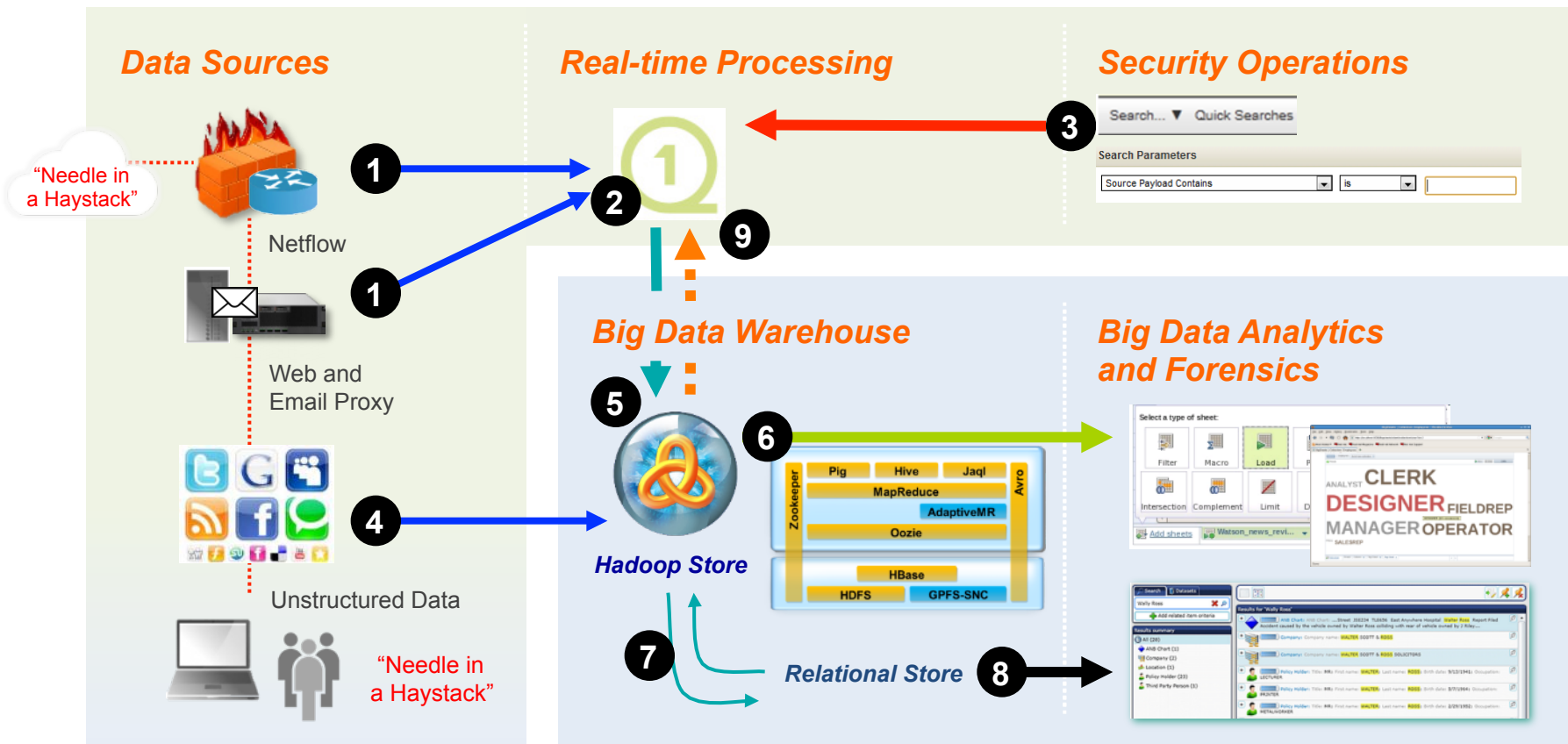
## Requirements

Source: proxy, email, unstructured text  
 Sample Size: >25GB /src  
 Query time: <45sec  
 Analytics: Multiple

## IBM Approach

1. Logs extracted, sent to QRadar
2. Event normalization
3. Custom filtered events sent to SOC
4. Unstructured data to BigInsights
5. Log / event forwarding to BigInsights
6. Custom BigSheets queries / analytics
7. Post-processed data storage
8. i2 link-based visuals / analytics
9. Update of QRadar real-time rule sets

# Use case #2 – Ad hoc query for specific data on multiple sources

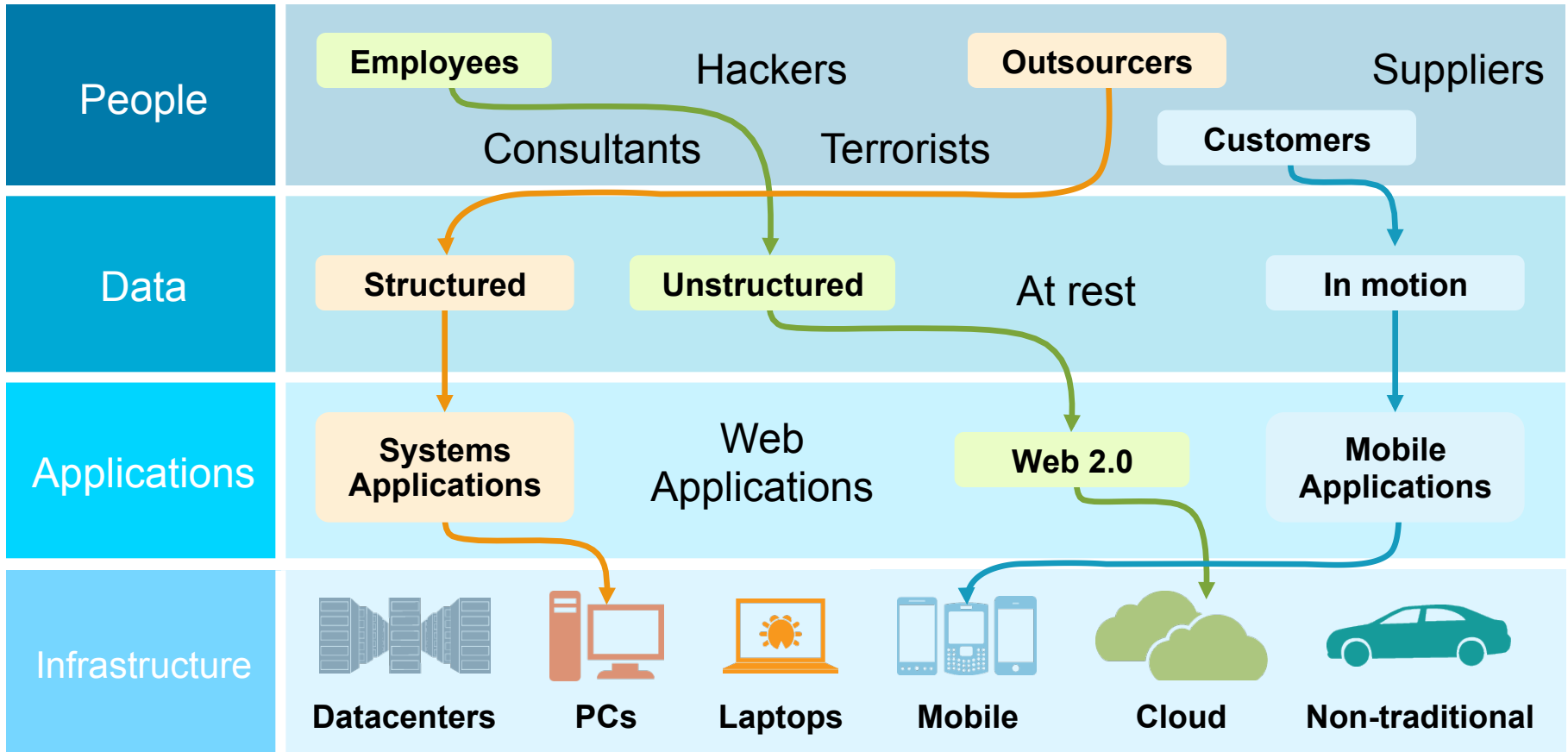


**Requirements**  
 Source: All  
 Sample Size: >20GB /src  
 Query time: <45sec  
 Analytics: Search for IP, FQDN and/or email address

- IBM Approach**
1. Netflow and logs sent to QRadar
  2. Event and flow processing
  3. Ad hoc payload search from SOC
  4. Unstructured data to BigInsights
  5. Events and flows sent to BigInsights
  6. Custom BigSheets queries / analytics
  7. Post-processed data storage
  8. i2 text-based, federated search
  9. Update of QRadar real-time rule sets



# Remove existing barriers between siloed detection mechanisms and integrate through Analytics



Attempting to protect the perimeter is not enough – siloed point products cannot adequately secure the enterprise

# Thank You

Sergio Mucciarelli  
Data Security Leader - IBM Italia  
[smucciarelli@it.ibm.com](mailto:smucciarelli@it.ibm.com)  
+39.335.1432985

---