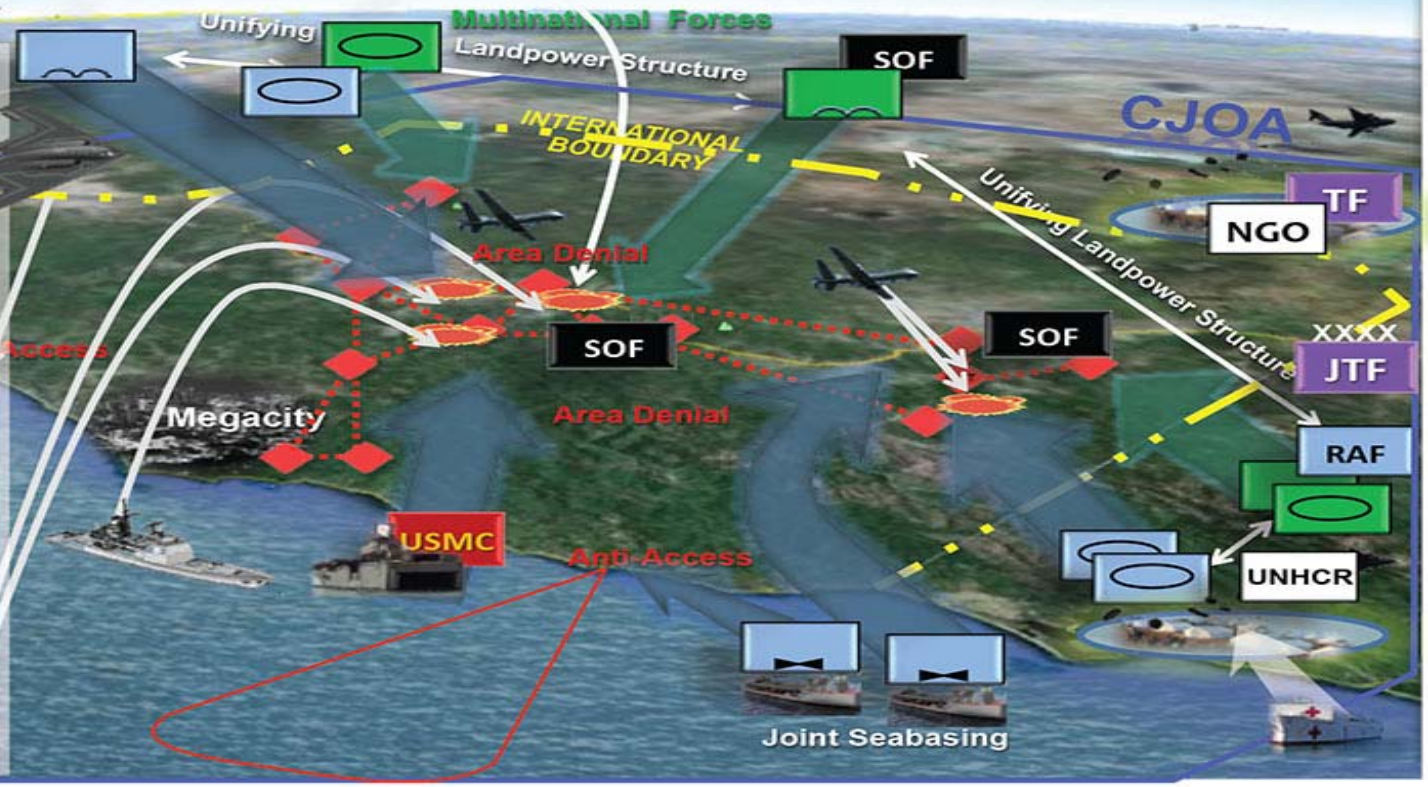
Talking Big Data

Dr. Robert C. Whetsel
Technical Director & Chief Data Architect
National Background Investigation Systems

Win in a Complex World

- Provide foundation for joint operations.
- Deploy and transition rapidly.
- Develop the situation in close contact.
- Maneuver from multiple locations and domains.
- Present multiple dilemmas to the enemy.
- Operate dispersed while maintaining mutual support.
- Integrate partners.
- Consolidate gains.

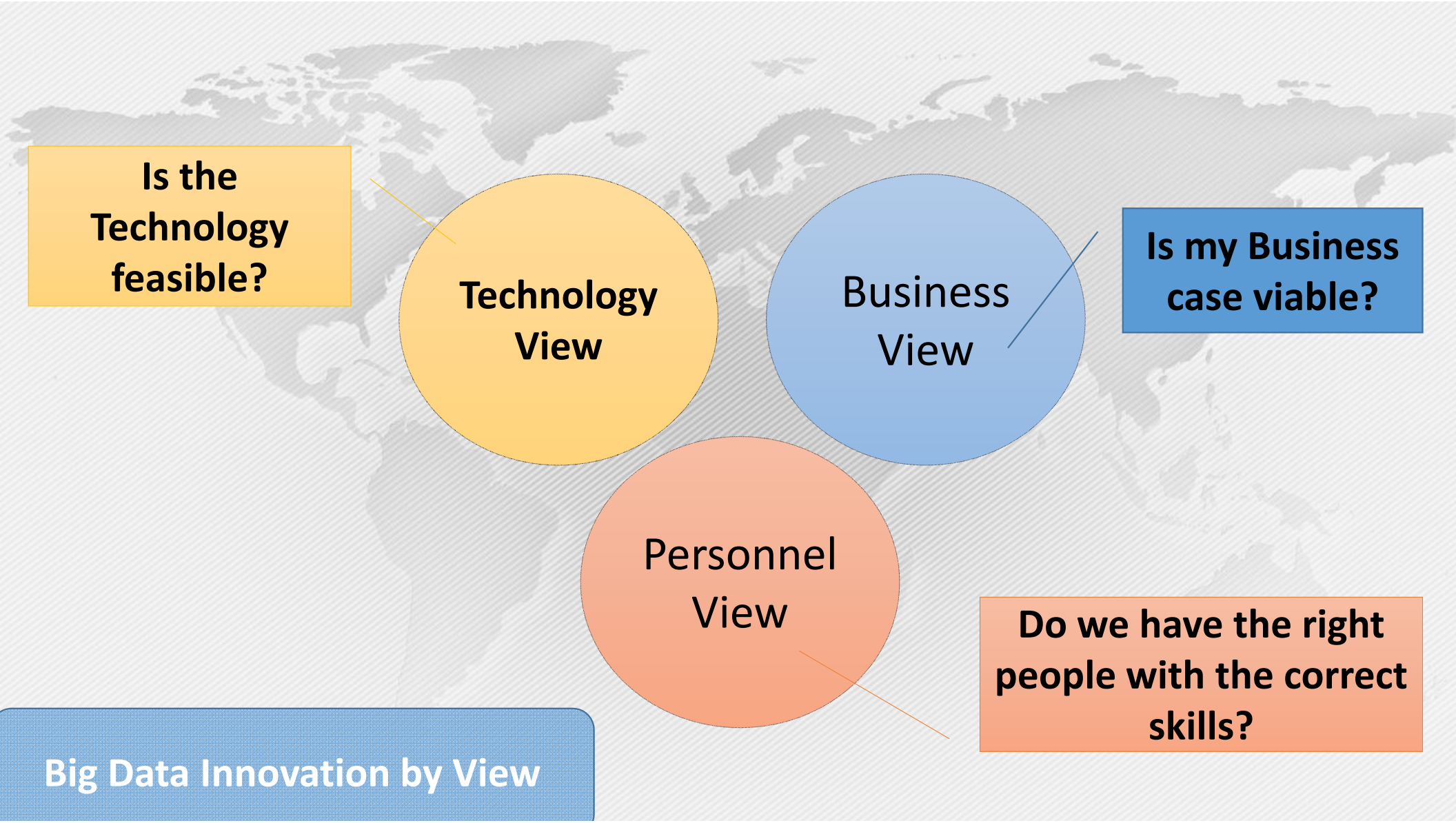


KEY

CJOA - Combined joint operations area
JTF - Joint task force
NGO - Nongovernmental organization
RAF - Regionally aligned forces
SOF - Special operations forces
TF - Task force
UNHCR - Office of the United Nations High Commissioner for Refugees
USMC - U.S. Marine Corps

Problem Statement

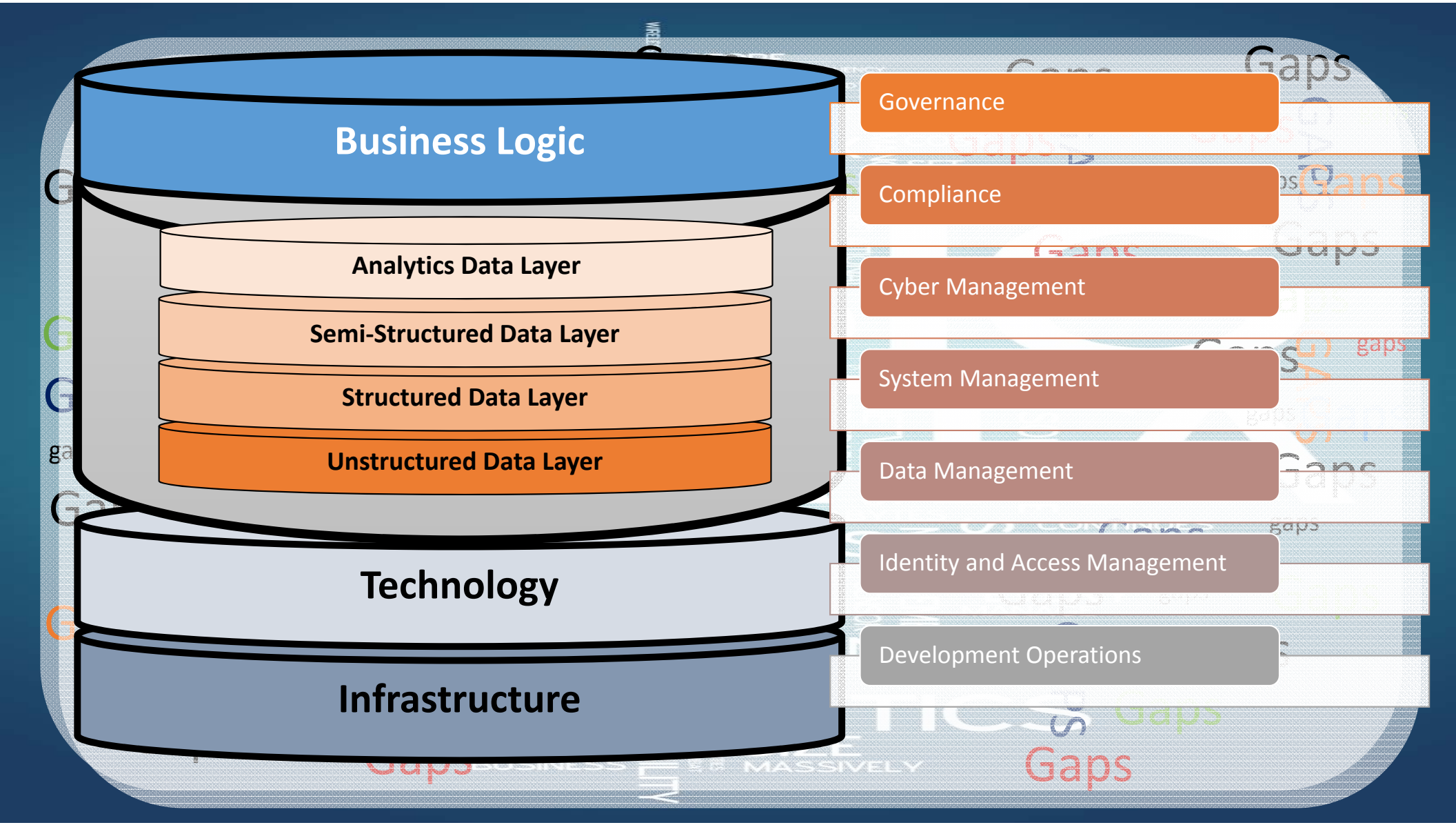
NBIS is a program directed to architect, design, build, secure, operate and maintain a replacement IT system for the Federal Investigative Service. This problem extends outside of the DoD and bleeds into the federal space. While I can speak to the DoD big problem set from a user's perspective, my problem set is much larger to include the Federal space's problem set. This adds an additional layer of complexity to what my team is trying to solve.

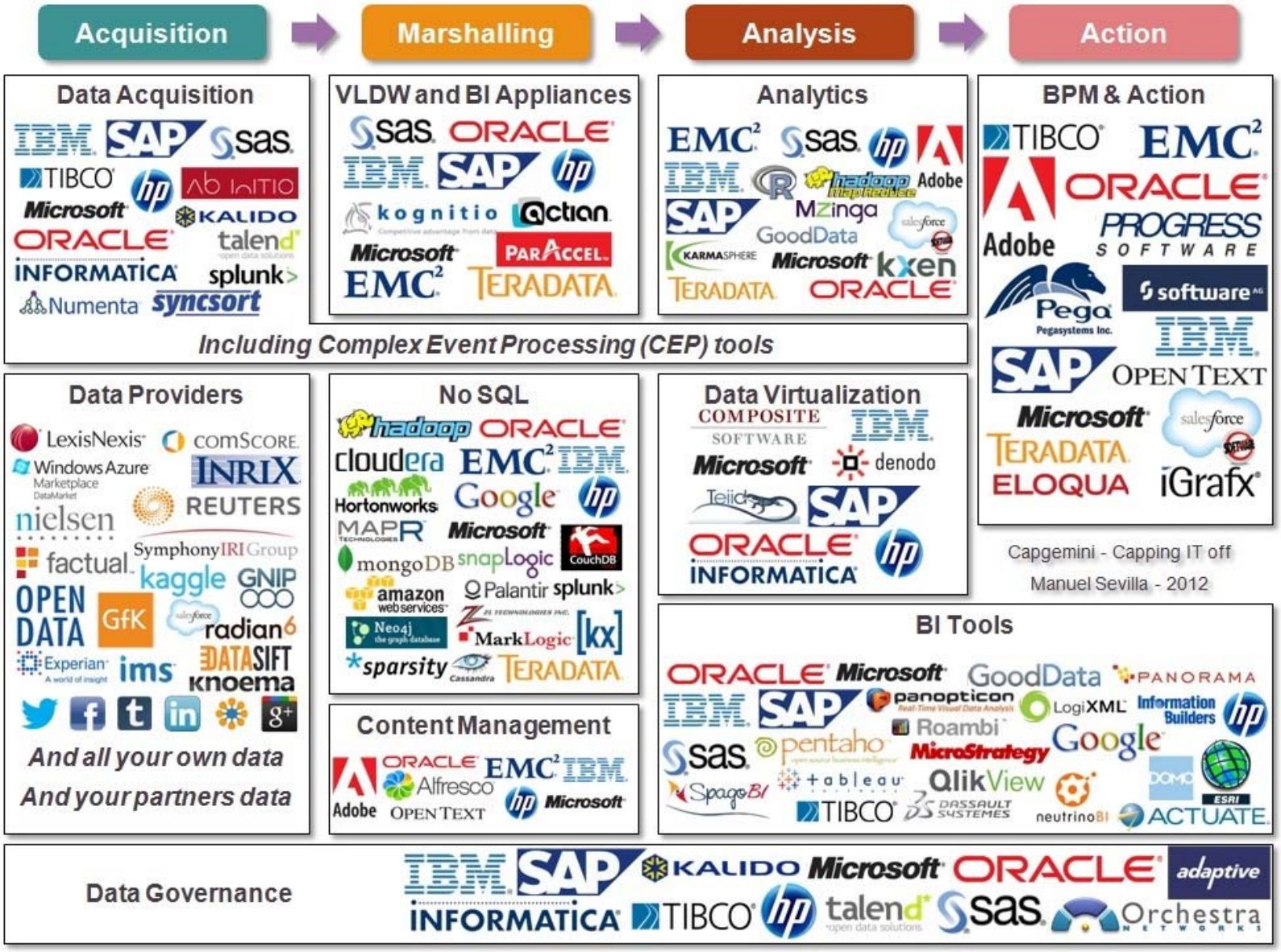

Is the Technology feasible?
Technology View
Business View
Is my Business case viable?
Personnel View
Do we have the right people with the correct skills?
Big Data Innovation by View

What Big Data Means to Me:

I have so much data that I cannot use traditional methodologies, software, hardware, or infrastructure to produce value in a reasonable amount of time.

Volume, Velocity, Variety, and Veracity


Business Logic
Analytics Data Layer
Semi-Structured Data Layer
Structured Data Layer
Unstructured Data Layer
Technology
Infrastructure
Governance
Compliance
Cyber Management
System Management
Data Management
Identity and Access Management
Development Operations



Capgemini - Capping IT off

Manuel Sevilla - 2012

Variety can be mathematically represented by approximation, the complexity of a big data set using Kolmogorov's Complexity Theory:

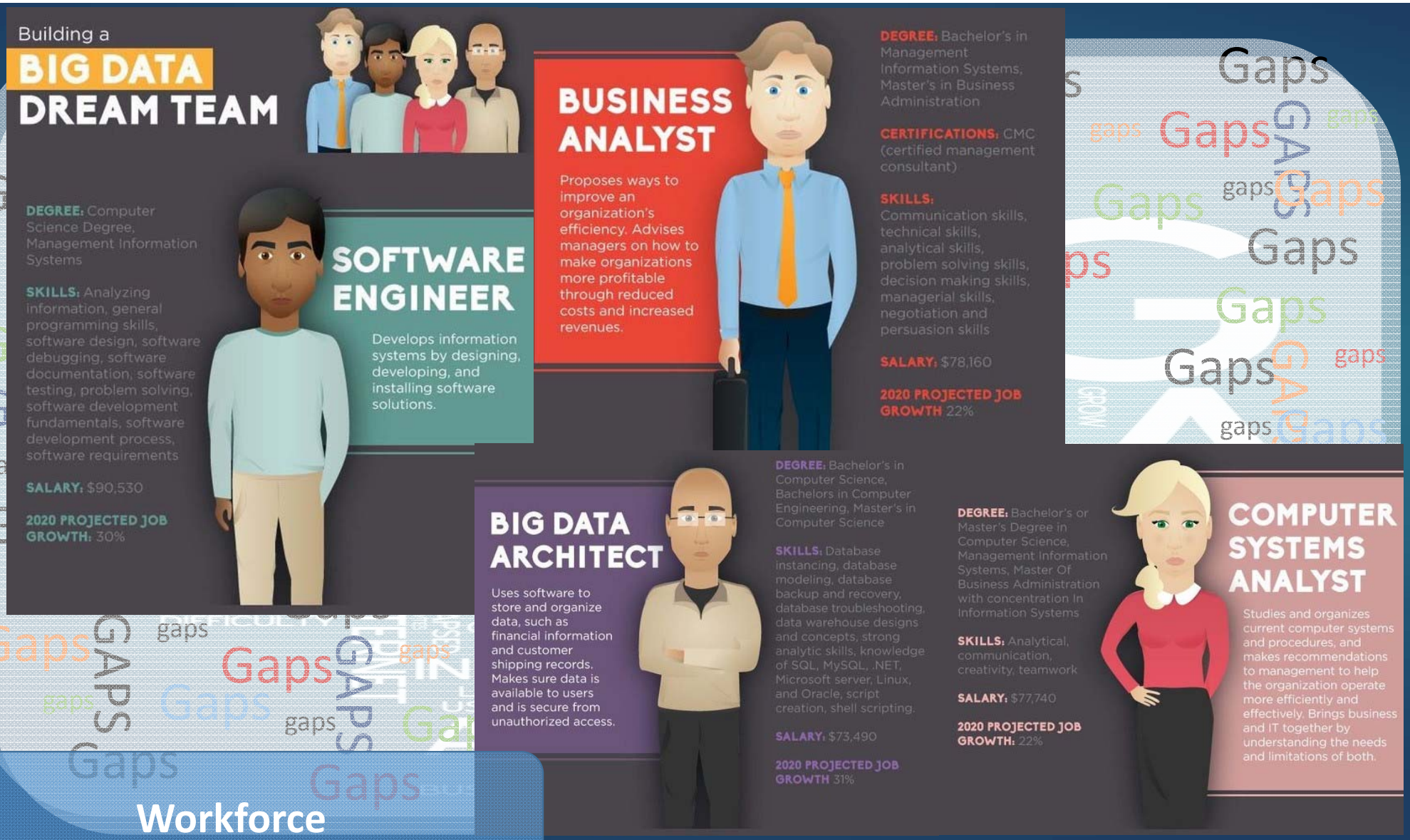
$$K = (|M, y|) \leq k$$

As extended by offering that complexity of Variety of *M* can be measured by the following formula:

$$Complexity(Variety\ of\ M) = \sum_{i=1}^{n} K(|M, y(r_i)|) \leq \sum_{i=1}^{n} k_i$$

MEASURING BIG DATA COMPLEXITY USING KOLMOGOROV'S COMPLEXITY THEORY (2016) *Robert C. Whetsel*

Words, Words, and Some Math…


Building a
BIG DATA
DREAM TEAM
SOFTWARE ENGINEER
Develops information systems by designing, developing, and installing software solutions.
DEGREE: Computer Science Degree, Management Information Systems
SKILLS: Analyzing information, general programming skills, software design, software debugging, software documentation, software testing, problem solving, software development fundamentals, software development process, software requirements
SALARY: $90,530
2020 PROJECTED JOB GROWTH: 30%
BUSINESS ANALYST
Proposes ways to improve an organization's efficiency. Advises managers on how to make organizations more profitable through reduced costs and increased revenues.
DEGREE: Bachelor's in Management Information Systems, Master's in Business Administration
CERTIFICATIONS: CMC (certified management consultant)
SKILLS: Communication skills, technical skills, analytical skills, problem solving skills, decision making skills, managerial skills, negotiation and persuasion skills
SALARY: $78,160
2020 PROJECTED JOB GROWTH 22%
BIG DATA ARCHITECT
Uses software to store and organize data, such as financial information and customer shipping records. Makes sure data is available to users and is secure from unauthorized access.
DEGREE: Bachelor's in Computer Science, Bachelors in Computer Engineering, Master's in Computer Science
SKILLS: Database instancing, database modeling, database backup and recovery, database troubleshooting, data warehouse designs and concepts, strong analytic skills, knowledge of SQL, MySQL, .NET, Microsoft server, Linux, and Oracle, script creation, shell scripting.
SALARY: $73,490
2020 PROJECTED JOB GROWTH 31%
COMPUTER SYSTEMS ANALYST
Studies and organizes current computer systems and procedures, and makes recommendations to management to help the organization operate more efficiently and effectively. Brings business and IT together by understanding the needs and limitations of both.
DEGREE: Bachelor's or Master's Degree in Computer Science, Management Information Systems, Master Of Business Administration with concentration In Information Systems
SKILLS: Analytical, communication, creativity, teamwork
SALARY: $77,740
2020 PROJECTED JOB GROWTH: 22%

Workforce

Data Science- Comprehensive knowledge of the state-of-the-art in data science and data exploitation techniques and methodologies, including data acquisition, cleansing, transformation, normalization, modeling, and visualization.

Data Engineering - Data engineer is premised on a skill set for developing reliable, scalable, completely automated data pipelines. That requires profound knowledge of every layer in the stack, beginning with cluster design, and spanning everything from Hadoop tuning to setting up the top chain responsible for processing data. Must be familiar with distributive computing, scaling, planning (design, build, test, production) and ingress/egress of data (batch, transaction and event-driven).

- Have I quantified my problem?
- Is my problem a real big data Problem?
- Where in the big data problem space does my data set fit?
- What resources are needed to solve my big data problem?
- Is my Problem Worth Solving?

Contact Me

Dr. Robert Whetsel
Technical Director
Chief Data Architect
National Background Investigation Systems
✉ robert.c.whetsel.civ@mail.mil
☎ 301.225.7517

http://robertwhetsel.com
www.linkedin.com/in/robertwhetsel/
✉rwhetsel@ravensong.com
☎Mobile: 240-344-2226