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
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

Is my Business case viable?

Business View

Do we have the right people with the correct skills?

Personnel View

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
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:

\[
\text{Complexity}(\text{Variety of } M) = \sum_{i=1}^{n} K(|M, y(r_i)|) \leq \sum_{i=1}^{n} k_i
\]
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, Bachelor's 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%
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?
Dr. Robert Whetsel
Technical Director
Chief Data Architect
National Background Investigation Systems
✉ robert.c.whetsel.civ@mail.mil
☎ 301.225.7517
www.linkedin.com/in/robertwhetsel/
✉ rwhetsel@ravensong.com
☎ Mobile: 240-344-2226