This Symposium brought to you by

www.ttcus.com

Linkedin/Group:
Technology Training Corporation
@Techtrain

Technology Training Corporation
www.ttcus.com
Big Data Analytics as a Service (BDAaaS)

Big Data for Defense & Intelligence Symposium
Rod Fontecilla, Ph.D.
Vice President
Agenda

• Big Data Analytics Overview
• Big Data Analytics Use Cases
• Methodology, Framework and Reference Architecture
• Q&A
Unisys Builds Large Complex Mission Critical Big Data Knowledge Repositories

On a Typical Day, DHS-CBP

- Processes 932,456 passengers and pedestrians.
- Processes 64,483 truck, rail, and sea containers.
- 470 refusals of entry at our ports of entry and 61 arrests of criminals at ports of entry.
- Seizes 13,717 pounds of drugs.
- Fusing multiple disparate data sources to provide useful information to CBP analysts.

*TASPO is THE program that provides protection against potential threats while facilitating commerce and people-flow through our nation’s border crossings. We process more than 1.3 Billion transactions a day.*

*Unisys has been supporting DHS for more than 15 years. Integrating more than 30 different applications all over the world*
The Big Data Analytics Problem

1. How Can We Ingest, Correlate, Entity-Extract, Entity-Map, Machine- and Human-Enrich the Never Ending Wave of Data?

2. While Providing a Variety of Search, Suggestions, Discovery Recommendations, and Visualization Mechanisms

3. And, Predict the Future to make better decisions
Unisys Differentiator – Building Enterprise Data Analytics Environment

Data Analytics

Data Volume

Low Volume, Variety, Velocity

High Volume, Variety, Velocity

Modeling and Forecasting
- Pattern Recognition
- Global Optimization
- Machine Learning
- Simulation

Business Intelligence & Data Warehousing
- SQL
- ETL
- STAR Schema
- Hadoop
- Map/Reduce
- Hive
- Cassandra
- HBase

Big Data & NoSQL
- Google
- BigTable
- Splunk
- Dynamo
- EMC
- MongoDB
- SAS & R

Leverage for large-scale analytics and data mining

Leverage for large-scale application development & information management

Low Volume, Variety, Velocity

High Volume, Variety, Velocity

Forward-looking (Predictive)

Backward-looking (Forensic)

Complexity

Scale-out

Integrate / Absorb

Extend

Multi-TB Turning Point

© 2013 Unisys Corporation. All rights reserved.
Building Analytic Environment

1. Work with business leaders and decision makers to understand and quantify data value chain
2. View data as an enterprise asset
3. Innovate through creation of new data products and services
4. Retrain staff and/or acquire Data Scientist skills
5. Integrate teams across big data, data warehousing, and business analysis
6. Revise information management strategies to incorporate big data
7. Develop new ways of capturing information e.g., mobile and streaming data
8. Identify and leverage previously unused internal and external data
Different domains need different insights
   - Essentially this is the art and science of taking a set of very complicated analytics and operationalizing or monetizing them such that they are consumable by customers.

They can manifest themselves as many things
   - Analytical "engines" running in a larger application (Amazon's recommender engine is a great Data Product)
   - Lists (e.g., Top 10 things I need to know today)
   - Entire applications (e.g., customer baseball cards)

However once they are defined, one thing is true for all
   - It takes a combination of domain agnostic analytic techniques together with domain specific knowledge to produce something relevant and consumable that can be monetized or operationalized.
Use Cases
Use Case – Federal Customer
Increase Market Share and Identify Customer Behavior

• Federal Customer was looking for guidance on leveraging their data to do predictive analytics to help drive their bottom line. Customer wants to leverage text mining analytics to link the unstructured data between their opportunities and potential offerings

• **Unisys Predictive Analytics Services:** The more money that a customer spends utilizing their services ultimately improves their overall revenue and profitability
  
  – **Results:**
    
    • Leverage transactional sales data to identify situations, indicators or actionable intelligence to increase business
    
    • Developed a statistical predictive model that identifies customers who have a high probability of decreasing their GSA spend in 2013. These customers would require additional actions to help maintain or increase their GSA spend. Additionally developed a machine learning algorithm to recommends similar services that vendors with should be qualified to support

• **Unisys Text Mining Services:** The ability to generate and understand opportunity leads will help customer focus their acquisition support services
  
  – **Results:**
    
    • Leverage opportunities from FedBiz Opps, and other reference and schedule data to create an indexing system that identifies high probability opportunity/schedule options to use as a lead generation tool
    
    • Customer can get relieve from manual analysis and assessment of opportunities by using this approach to triage opportunities in an automated way allowing analysts to focus acquisition assistance on specific opportunities
• **Challenge:** Large intelligence customer needs to understand big data use cases and how it is changing the landscape from data to insights

• **Solution:** Prepared curriculum on Big Data and Data Science including current state of the art and new techniques
  – Related use cases and approaches of known commercial entities to the “protective services” domain

• **Result:** Multi-course training program on existing contract to cover:
  • Big Data Introduction
  • Data Mining and Information Retrieval
  • Social Network Analysis and other special topics
  • Data exploitation techniques and mitigations
Use Case – Commercial Clients
Analyze IT Service Management to Increase Operational Efficiencies

• GMS is Unisys’ Global Manage Services and maintains a rich database of details of the incidents, tasks, chats and customers surveys related to each transaction

• GMS challenges relate to analyzing their full data and leveraging this to help improve their overall processes. This includes lack of ability to view entire time frames, provide enhanced reporting, understanding customer sentiment, automating manual processes and predictive analytics

• **Unisys Process Automation & Reporting Services:** GMS’s current reporting process involves a number of manual processes to calculate variables, trend data and populate reports required for client review
  
  – **Goal:** Leverage our big data environment (including data ingestion and visualization) and the appropriate machine learning algorithms to streamline and improve the reporting process
  
  – **Results:** Recreated GMS operational reports within our big data environment including all variables, calculations and graphics. Currently developing machine learning algorithm to populate manually calculated variables

• **Unisys Sentiment Analysis Services:** GMS provides surveys to their customers in order to obtain feedback on overall service and understanding the unstructured text can help improve overall customer service
  
  – **Goal:** Analyze the unstructured comments text to identify the customer sentiment for the entire history of the database (5+ years, 1 million records)
  
  – **Results:** Analyzed all Additional Comments unstructured text, identified key terms that lead to both positive and negative experiences, determined sentiment score for each customer. The sentiment scores were plotted for each customer for the years 2009 to 2013 by customer.
This shows the average sentiment of the clients over time by Quarter. Provides opportunities to business leaders to address issues and document best practices.
Unisys Big Data Analytics Methodology, Framework and Reference Architecture
Big Data Analytics Service Areas

- Master Data Management
- Storage Strategy
- Data Collection/Analytics
- Content Normalization
- Text Analytics
- NoSQL Database
- Hadoop
- ETL
- RDBMS
- Knowledge Discovery

- Enterprise Data Warehouse
- Visualization Tools
- Link Analysis
- Business Intelligence
- Geospatial Services
- Mobile Applications
- Identity Management
- Pattern Recognition
- Predictive Analytics
Modeling Components

- Decision Making & Forecasting
  - Provide actionable intelligence into the future state

Models

- Statistical model applied to input data that separates the portion of volume due to each of the variables or factors. We use the term model, because it is a simplification of reality.

Data

- Internal Data
- Demographic Data
- 3rd Party Data
Unisys Differentiator – Data Analytics Methodology and Framework

1. Ingest Multiple Sources
2. Manual Discovery & Understanding
3. Use R to create multiple views from tables
4. Load multiple views into Hadoop for analysis
5. Stitch all views together to create de-normalized columnar tables
6. Validate de-normalized columnar table definitions
7. Refine and Improve
8. Decision Point
9. Decision Point
10. Feedback

- BI Analytics
- Dashboards
- ETL Analysis
- Data Warehouse
- Using tools such as Tableau, Splunk, etc

- Check Models
- Apply Linear Regression and Clustering
- Define Variables
- Create Models
- Predictive Analytics
- Statistical Analysis
- Using tools such as R, SAS

- Monitor, Measure, Analyze and Improve
- Create Data Products into Production
- - Implement Real Time into Production
  - - Update Infrastructure
  - - Integrate with other applications

© 2013 Unisys Corporation. All rights reserved.
Data Products Become the Drivers to Identify new Revenue Channels, Cost Savings and Increase Efficiencies

Your Customers

Feedback

New Revenue Channels

Better Service Cost Savings

Data Analytics Environment Knowledge Repository

Populate

Internal Data Sets

Analytics Engine

External Data Sets
Unisys Data Analytics as a Service

Data Ingestion and Cleansing
- Ingest Data
- Assess Data Quality
- Cleanse Data
- Standardize Data
- Enrich Data

Unisys Data Analytics as a Service
- Dashboards
- Business Intelligence
- Predictive Analytics
- Machine Learning
- Text Mining
- Pattern Recognition

Clients Real Time Dashboard
- Real Time view
- Visualization artifacts
- Export/Import utilities
- Ad-hoc reports
- Pre-defined reports

Unisys Data Products
- Sentiment Analysis
- Recommender engine
- Unstructured documents integration
- Resolver analysis
- Event logs analysis and monitoring

Running Both At Unisys and At Amazon

Customer ability to access and use reporting capabilities

Customer ability to use Data Products and create new ones
The Unisys Data Analytics As A Service Environment – Ready for Business

Ingest Data
- Data is assessed, cleansed and organized
- Partitioning for performance

Normalize
- Transformation of data at scale
- Integration of algorithms, taxonomies and/or 3rd party data

Productize
- Construction of consumable data products including:
  - Indexes
  - Recommenders
  - Correlated Entities
  - Aggregations

Running on Unisys Infrastructure and at Amazon
Unisys Data Analytics as a Service

Your Cloud | Our Cloud | Public Cloud

Unisys Data Analytics Roadmap

Engagement Conceptualization
Brainstorming and Roadmap

Data Acquisition
Sources from Intelligent Sensors to Social Networks

Data Transformation
Data Refinement and Global Views

Data Analytics
Smart Computing & Intelligent Analytics

Data Productization
Operationalization or Monetization of Analytics

Idea generation,
Linking the dots across organizations,
Data Rationalization,
Deliverables Definition

Data Management Strategy,
Data Consolidation,
Data Security,
Storage Rationalization

Information Extraction,
Consolidated Views,
Database Consolidation,
Data Enrichment,
Data Protection

Ad-hoc Analysis,
Pre-computations and Aggregations,
Data Mining,
Machine Learning,
Predictive Modeling

Dashboards & Reports,
Forecasting & Modeling,
Search & Retrieval

Unisys help you take a strategic approach to Big Data that covers the full lifecycle from data acquisition to analytics - whilst leveraging your existing assets
Thank you