Making Your Data Work for You
By James Bessin, Vice President, Xerox Federal Solutions
The digital universe is growing at an astounding rate and expected to double every two years, according to research firm IDC. The explosion and inherent value of data, coupled with expanded availability of technology to analyze, synthesize, and visualize data, poses both challenges and opportunities for Federal agencies.

The questions are no longer “if” and “why” agencies are using data analytics, but “how.” To gain meaningful intelligence from data, agencies must mine, cleanse, tag, and aggregate their data to transform it into usable information to improve service; maximize efficiency; and reduce fraud, waste, and abuse.

As a global leader in business process outsourcing, Xerox understands how to operate at scale. For years, we’ve focused research and development efforts on finding new and better ways to work with information and to realize the benefits from the inherent value of data to increase service quality, reduce costs, and improve program integrity.

This paper presents practical ways we use data analytics to enhance agency efforts to improve operation.

Federal Data Analytics in Action

Data analytics is not a single technology, but an approach to harnessing data integration, information extraction, visualization and predictive analytic tools in order to improve how Federal agencies achieve their missions. Below are just a few examples of how we’ve woven data analytics into our solutions.

Document Management

Nearly every Federal agency is plagued by a daily flood of both inbound paper and volumes of new electronic records. Rather than working to tread water in a sea of documents, agencies need to make their documents work for them! We process more than 1 million documents daily and integrate them into real time digital workflows.

Many documents are containers for unstructured data and on their own are not very smart. Historically, automatically extracting knowledge from unstructured text is difficult due to multiple document standards, complicated document structure, and writing style variation. One approach we have taken is to bring natural language processing to the document to make it easier to extract the relevant information using features such as automatic indexing security and accessibility. Other documents don’t contain so much unstructured text but more closely resemble a form. In those cases, we’ve used an Automated Document Recognition (ADR) solution to automate the recognition of forms and extraction of field level data.

Clustering technology classifies documents by splitting them into smaller groups of similar objects whose characteristics of similarity may not be known in advance. An example of cluster analysis is looking at fraudulent documents and finding other documents with similar linguistic features. Similarly, we use clustering technology in the Federal digital mailroom to route similar documents to the same work queue.

Classification identifies the categories in which new data points belong based on prior experience. We use advanced software to analyze and filter all types of documents, such as letters, forms and invoices, to determine what type (category) of document they are, or to whom they belong, in a business process.
Customer Care

Not only can data analytics technology help Federal Agencies uncover valuable insight on their intellectual property; it also gives Federal employees on-demand access to information to serve citizens better.

Xerox has more than 1,000 customer service representatives via Customer Care centers, handling more than 20 million calls annually in support of Federal Agencies.

With more information at their fingertips, customer service representatives can reference past interactions and pull data about a citizen to quickly and effectively solve issues, ultimately increasing customer service and experience.

Each year, we field approximately one billion public sector service calls, and our call center managers and agents actively use analytics-based dashboards and call support in support of their clients’ missions. Incorporating functions like predictive analytics, natural language processing, and machine learning along with key word identification, gives structure to the huge volume of call traffic and automates the delivery of more accurate insight into customer sentiment and operational metrics. To do this, we use a diversity of data such as:

- Call scripts and notes
- Agents training and payroll
- Switch data and device data
- Citizen surveys
- Caller demographics
- Knowledge bases
- Demand history, Interactive Voice Response, emails
- Social networks

Through predictive and prescriptive analytics built on a machine learning platform, we’re developing technology designed to predict and resolve customer needs before they are made. Machine learning continually analyzes customer care data, learning how to solve customer queries in the same way a live care agent would. Rich analytics capture the context of every customer interaction across multiple direct and social channels to provide detailed insights into customer behavior, problems, and solution success.

We also offer an intelligent self-service tool that automates interactions between an agency and its constituents. Self-service care learns from every interaction and detects how human agents diagnose customer problems and offer solutions. In doing this, it quickly develops the intelligence it needs to understand and solve queries itself, without having to be programmed.

Workforce Management

Between budget constraints, an aging (largely retirement-eligible) workforce, and heightened citizen expectations, agencies need to operate at maximum efficiency. Data on workforce processes and procedures can help agencies identify bottlenecks that are slowing down employee productivity.

This data allows agencies to uncover problems and develop solutions with technology to streamline their workflow processes. Employees’ time can then be diverted to more meaningful mission-related work tasks. Built-in analytic tools measure and collect data, such as clicks and average visit duration, providing valuable feedback in developing future content.
In addition to bolstering productivity and reducing Federal workers’ more mundane tasks, data will play an increasingly central role in training and development to prepare the next generation of the Federal workforce.

**Human Resources leaders will be able to leverage data analytics to measure the true business impact of learning activities, rather than simply confirm that required coursework was fulfilled.**

Agencies will be able to use predictive analytics and diagnostics to determine whether employees are performing as expected after completing learning interventions, and this data will be used to inform the selection of future learning activities.

## Best Practices in Data Analytics

To support data analytics projects, we have adapted our program management methodology to provide managers with an integrated set of best practices for deploying data analytic systems, integrating data sources, refining data analysis, transferring knowledge, and receiving feedback. This is a critical component of our approach to project management, quality management, and data governance allowing us to manage and execute high-quality technical solutions both on time and within budget. Best practices incorporated in our data analytics methodology include:

**Authentic Testing:** Thorough testing of content analysis including application to real problems by subject matter experts who understand the terminologies (semantics), rules of behavior, and decision rules in a specific domain.

**Updates and Improvements:** Continual refinement based upon industry trends, feedback from data analytic projects, surveys and evaluations, and internal improvement efforts.

**Standardization with Customization:** Assurance of alignment with important industry standards while specifically addressing individual agency needs.

**Text Categorization:** Linguistic analysis technologies and machine learning algorithms using the data mining techniques of clustering and classification to tag text by subject or attribute.

**Image Categorization:** A classification engine capable of categorizing multiple everyday images content types, including buildings, airplanes, books, faces, and documents. The result includes both a written description of the image and image tags by subject or attribute.

**Fact Extraction:** Determination of the meaning of statements and the relations that link them together.

**Advanced Text Analysis:** Linguistic engines that recognize more advanced elements or concepts such as dates, addresses, company and person names, locations or organizations, and even syntax and semantics of the text. The resulting analysis is a knowledge search, rather than a keyword search, that:

- Understands the content of both structured and unstructured data
- Takes into account the context within the document
- Tags specific facts for future use
- Finds relationships and uncovers smart insights among related facts
- Summarizes content and allows analysts to focus on the information that is pertinent to their business or oversight processes
Hadoop and Spark: Open-source software frameworks that provide massive key-value storage, hold enormous processing power, and have the ability to handle virtually limitless concurrent tasks or jobs.

Graph Analytics: A high-performance, in-memory graph analysis engine that exploits parallelized search and compact graph representations that are amendable for traversal and other queries for which key-value processing is inefficient.

Predictive Modeling: Use of available static or streaming information, machine learning, experiment design, and a wide array of algorithms to generate actionable predictions.

How Can Data Analytics Help You?

All Federal agencies today share a common goal to cut costs and increase efficiencies, and we have the answer with our data-driven approach to your business processes. We have a solid understanding of government operations and a vision of emerging federal government service trends.

Let us help your agency address your business imperatives, substantially reduce cost, improve service levels, and direct valuable resources to initiatives that generate customer satisfaction.

www.xerox.com/govsolutions