Organization, Education & Development for Data Analytics Capabilities

A WHITEPAPER FROM THE AEIC DATA ANALYTICS COUNCIL
Acknowledgements

The Association of Edison Illuminating Companies would like to acknowledge the contributions of the following members of the AEIC Data Analytics Council who produced this white paper:

Josh Power, Oncor  Team Leader
Diana Mahoney, Eversource
Denise Ralston, First Energy
Loretta Townsley, FPL
Matthew Tovar, Oncor
Tom Siudak, First Energy
Contents

Introduction .............................................................................................................................................. 5
What Skills Are Needed? .......................................................................................................................... 5
  Data Scientist, Data Engineer, Data Analyst .......................................................................................... 5
Skill Development Options - Education .................................................................................................... 6
  Data Scientists ..................................................................................................................................... 6
  Data Engineer ...................................................................................................................................... 7
  Data Analyst ...................................................................................................................................... 7
  Data Analytics Curriculum for Everyone ............................................................................................. 7
  Curating a Data Analytics Curriculum .................................................................................................. 8
  Data Analytics Curriculum Example ..................................................................................................... 9
Online Courses ....................................................................................................................................... 10
On Campus Coursework/Classes ............................................................................................................ 10
Onsite Classroom – Bring in Educators/Experts ...................................................................................... 11
Onsite Classroom – Internal SMEs .......................................................................................................... 11
Conferences & Industry Associations ...................................................................................................... 11
Skill Development Options .................................................................................................................... 11
  Community of Practice/Grass Roots Effort ........................................................................................... 11
  Center of Excellence (CoE)/Analytics Community ............................................................................... 12
  Interns .................................................................................................................................................. 12
  Job Rotational Programs ...................................................................................................................... 13
  Partner Cross Functional Teams (IT & Business) .................................................................................. 14
Setting Expectations from Educational Investments ................................................................................ 14
  Expectations of Data Analytics Resources ........................................................................................... 14
  Measuring Results ............................................................................................................................... 14
Change Management ............................................................................................................................... 14
  Gaining Executive Support .................................................................................................................. 14
  Leveraging Diversity ............................................................................................................................ 15
  Educating the Organization about Data Analytics (Create Awareness) ................................................ 15
  Communication ................................................................................................................................... 16
  Building a Roadmap ............................................................................................................................. 16
References ............................................................................................................................................... 17

AEIC DATA ANALYTICS COUNCIL
Introduction

Purpose
The purpose of this whitepaper is to provide an overview of strategies to build Data Analytics capabilities within a utility business.

Need/value
Data Analytics is a growing field in all industries. As data is more readily available due to the insertion of smart devices across many business units, and the increase in availability of tools to harness the plethora of collected data, utilities can realize great value from the insights provided through Data Analytics. To capitalize on the value from Data Analytics, the skillsets to manage large amounts of data, perform analytics to develop meaningful algorithms, and visualize the results is critical for success.

Overview – Need for a Roadmap
This paper will cover topics including a description of the desired Data Analytics related skillsets, a variety of means to acquire said skillsets, organizational structures to consider, performance measurement of a training program, and change management. There are many topics to consider in determining the right solution for each Utility. Once the options are reviewed, a roadmap will be needed to plan the steps required to achieve the desired goal of enabling Data Analytics capabilities.

What Skills Are Needed?
There are basic job classifications and skill sets required to progress Data Analytics (DA) strategies within an organization. This section describes the job classifications and the basic skillsets required in specific DA areas. Additionally, this section will cover a wide variety of skills of value in managing large amounts of data, analyzing data in a meaningful way, and visualizing results for the organization to put to use.

Data Scientist, Data Engineer, Data Analyst

a. Data Scientist
   A Data Scientist has skills in transforming and combining multiple raw datasets as inputs to machine learning algorithms. They have a strong background in statistical models and machine learning, and are typically skilled at programming in R, Python, SAS, and/or Matlab. They can manipulate large amounts of data and create predictive and prescriptive models. A Data Scientist is also adept at telling the story based on data insights through visualization.

b. Data Engineer
   The Data Engineer develops, constructs, tests, and maintains architectures (databases and large-scale processing systems.) The Data Engineer recommends ways to improve data reliability, efficiency, and quality using a variety of languages and tools to find the best way to collect data from systems so that the information can be used by the Data Scientist and Analyst.
effectively. The Data Engineer also develops data set processes for data modelling, mining, and production.

c. Data Analyst

A **Data Analyst** is adept at querying structured data to perform analyses to answer business questions. A Data Analyst has a solid understanding of the business subject area and the skills to extract and visualize insights from a single source of data. They have basic programming skills in transforming and combining raw data. A Data Analyst should be highly skilled in SQL to query and transform data. They are also adept in the use of business intelligence (BI) tools to visualize data to answer ad-hoc questions, create reports, or create dashboards.

**Skill Development Options - Education**

Given the immense business value of building strong Data Analytics capabilities within your organization, what is the best way to develop these skills?

Obviously one of the best ways is through Education. However, there are so many educational opportunities from which to choose, it can be a daunting task for organizations (and employees) to choose the best ones to fit their learning needs.

It is recommended to start with the skills needed for the three specific roles highlighted above: Data Scientists, Data Engineers, and Data Analysts. We will discuss the skills and education needed for each below.

**Data Scientists**

This is the most specialized and technical of the three roles. This role requires years of education and hands-on experience to become a professional Data Scientist. One recent [KDNuggets](https://www.kdnuggets.com) blog reported that currently around 88% of Data Scientists have Masters Degrees and almost 45% have PhDs. Internal employees with Math, Statistics, Computer Science or Engineering degrees are all good candidates to become Data Scientists. For those who have the passion to pursue this path, we recommend Education Assistance programs that are designed to help fund advanced degrees for individual employees. This, of course, will take time and perseverance.

Fortunately, not everyone needs to become a Data Scientist. Instead of grooming Data Scientists internally, most organizations choose to hire (or consult with) a few Data Science professionals, who then partner with **Data Engineers** and **Data Analysts** to drive data analytic capabilities within their organizations.

While Data Scientists typically come on board with strong technical expertise, development opportunities still exist. Data Scientist need to possess strong Business Acumen, Communication and Teamwork skills. These “soft skill” training programs already exist in most organizations and can be leveraged to attract, grow and retain top Data Science talent.
**Data Engineer**

Data Engineering skills are also very technical, involving specific Information Technology (IT) related expertise in areas such as programming, data architecture, data administration and/or data security. In most cases, these skills are initially developed in college or university degree programs. Organizations usually hire (or outsource) this expertise, then provide continuing education to strengthen and broaden the existing IT skill base.

In the past, since the Data Engineering skills were primarily found in the Information Technology (IT) department, technical training to develop and enhance those skills also resided within IT. However, as an organization builds its data analytics capability, IT technical skills will become more important to a much broader cross-section of the organization. Below, we will discuss ways to leverage existing expertise and develop new programs to expand IT education outside the IT function.

**Data Analyst**

The Data Analyst role is the conduit between the IT function and business operations. It is the role that traditionally takes data from IT, then performs analytics on that data to solve business problems.

With the recent “democratization of data” making more information available to more people in the organization, the role of Data Analyst has expanded to include anyone in the organization who uses data to make business decisions. Therefore, education to build the Data Analyst skill set will need to be available to more employees across the organization. In fact, a September 2019 Harvard Business Review article, in conjunction with McKinsey Analytics, suggests that “Analytics training will become a standard part of all employee development.”

**Data Analytics Curriculum for Everyone**

So how do you build a data analytics curriculum for everyone that will support the need for Data Engineering skills outside of IT and the need for Data Analyst skills throughout the entire organization? With such a broad and diverse audience, we suggest a tiered approach that can be customized to meet individual employee needs.

In a data-driven culture, it is important for everyone to understand basic analytical concepts. An organization can build a strong analytics capability by providing Foundation—level education that is understandable, accessible and encouraged for everyone. From there, employees can choose their desired levels of expertise from Intermediate to Advanced analytical capabilities based on their job functions and their own professional aspirations. Expected behaviors at each level include:

- **Foundation**: Understands basic analytical concepts and the value of data to improve processes and make business decisions.
- **Intermediate**: Possesses a firm grasp of data science foundations and problem solving methodologies. Understands the value of Artificial Intelligence (AI)/ Machine Learning (ML) and
is able to support various data science-driven projects and communicate their value to the business.

- **Advanced**: Demonstrates a high level of competence in all aspects of data science and advanced problem solving. These are budding data scientist with strong analytical capabilities, project leadership and communication skills.

As there are different levels of analytics expertise, there are also many different skills needed to build strong analytical capabilities. We consolidated these skill sets into six major categories: Project Leadership, Problem Solving, Statistics, Data Visualization, Data Science and Cloud Services. These categories are not all-inclusive but provide a good base from which to start. Other disciplines, such as Commercial Skills and Financial Acumen should be considered as well.

Building a customizable Data Analytics curriculum allows for diversity of levels and skills throughout the organization. For example:

- Individuals with advanced-level Statistics and Data Science skills may want to develop Foundation or Intermediate – level skills in Problems Solving and Project Leadership to work more effectively with cross-functional operations teams.
- On the other hand, individuals with strong Problem Solving and Project Leadership skills may want to broaden their Data Visualization and Data Science expertise to solve more complex business problems.

As individual employees broaden their skills in more disciplines at higher levels, the data analytics capability of the entire organization grows.

**Curating a Data Analytics Curriculum**

Now that skills and levels are more clearly understood, it is time to find the right education and training to fill-in the skill gaps. Education is everywhere and most employees know how to find answers quickly. Ask anyone how to create a drop down menu in Excel or define “artificial Intelligence” and they will probably be able to find an answer.

The problem is more a question of quantity and quality of the education provided and consumed. This is where the Data Analytics Council might be able to help. In our collective efforts to improve the Data Analytics capabilities of our respective organizations, we have looked at hundreds of potential training options.

We have also researched and benchmarked the Data Analytics training capabilities of other companies outside of our industry. One thing we found immediately is that most companies, regardless of industry, are not much further along than utilities are. Most are just starting to figure out the best way to build “big data” analytics skills throughout their organizations. It appears that no one company has created the perfect approach.

We have also found there are many high-quality, low-cost training options. Here are some examples of the best, most cost-effective educational resources we found by category:
• **Project Leadership:** Leveraged existing in-house courses, Project Management Institute and Cprime / Blue Agility
• **Problem Solving Methods and Tools:** Leveraged in-house programs such as Lean and Six Sigma for traditional problem-solving, CRISP-DM white papers for Big Data analytics
• **Statistics:** Kahn Academy
• **Data Visualization:** YouTube, In-house Power BI or Tableau “Super Users”
• **Data Science:** YouTube, EdX, Udacity, Data Camp, Coursera, and In-house Data Scientists
• **Cloud Services:** In-house IT Architecture team, Pluralsight, Azure, AWS or other platform providers

**Data Analytics Curriculum Example**

In exhibit 1 below, you will see an example of how one company put together a data analytics curriculum to strengthen the skills within their company. This utility created a customizable scorecard for employees to check off the learning activities in each category at their appropriate skill level. Many employees had already completed some of these courses and just needed to fill in the gaps.

You Tube videos were often used. To improve tracking, these links were packaged with a few post-video Knowledge Check questions to assess comprehension. Tools such as Lectora, Adobe Captivate and Articulate 360 were used to track completions in the Learning Management System.

To increase employee awareness and participation, pins were awarded to those employees who completed the different levels.
Online Courses
As mentioned in the above section on Curating a Data Analytics Curriculum, Coursera, Udacity, EdX, Khan Academy, Code Academy, and many other on-line education options are available for little or no cost. If your utility has a formal learning and development department, chances are some of these online education options are already in use.

On Campus Coursework/Classes
Advanced Analytics courses are available for enrollment in most universities these days. Universities typically hold these classes within colleges of Business (Business Analytics) and/or Engineering or Science (within the Computer Science curriculum.) Prerequisite courses in statistics, math, and software development can often be taken at local community colleges.

Texas A&M, Arizona State, Georgia Tech, The University of Oklahoma, and other top-ranked universities offer Masters of Analytics degree programs on-line for less than $10,000. These programs provide good value if your utility offers continuing education assistance to its employees.

Exhibit 1 – Data Analytics Curriculum Example
Onsite Classroom – Bring in Educators/Experts

Bringing in a professor from a local college or university is oftentimes an option for in-person instructor led training. Class topics, duration, and number of students is often negotiable. Reach out to the Analytics department head at your local university for more information.

Onsite Classroom – Internal SMEs

Encourage the Data Analytics resources on currently staff to train other resources. This is a good way to augment data and analytics skills of your staff with deep functional knowledge. Encourage informal “brown bag” sessions on Analytics topics. Encourage knowledge sharing sessions at multiple levels of experience – from beginners through advanced techniques.

Conferences & Industry Associations

Conferences and Industry associations may provide educational opportunities alongside their meetings. It is worth noting these educational opportunities tend be conducted by a vendor or service provider and therefor may be biased towards an approach or product suite.

Skill Development Options

Community of Practice/Grass Roots Effort

A community of practice is a group of people who share a concern or a passion for something they do and learn how to do it better as they interact regularly.

A great way to build and keep skillsets fresh within a company is to create a grass roots group for employees with advanced data and analytical skills. In large organizations, many employees only get to work with others on their team or in their department. Community of practice groups will open networking opportunities across the company with others they may never interact with. Executive sponsorship is important and shows the members that they have the support of the company and their leadership, but the group should be organized and run by members of the group. The community’s leadership should understand the goals of the group and how to ensure the benefits are important to its members. A key thing to remember is that the members choose the direction of the group, not the company. While the group may be comprised of employees with advanced skills, the group should also be open to employees who may not have the skills today but are interested in learning and developing these skills. Successful groups will include diverse members from different areas of the company.

Meetings should occur on a regular basis and have of mix of topics that focus on the use of data and advanced analytics. While many of the meetings will be members sharing presentations of their projects, special guest speakers (consultants/vendors) from outside the company can keep the group informed on the latest technology and trends. No matter what the topic, the key is to encourage participation and networking.

In addition to meetings and speakers, special events throughout the year can enhance the participation and engagement from members. One example is a hackathon. A hackathon is a short event (usually a day or two) that groups members into small teams to develop and present solutions to a distinct
business problem. Having internal hackathons creates the opportunity for innovation, exposure to different tools, promote knowledge sharing and showcases the best talent in the company. Some utility hackathon examples: analyzing safety data for patterns and trends, how AMI data can benefit customers, how to lower costs with the better utilization of grid assets. If nothing else, a hackathon will help energize the group and encourage them to think outside of the box.

Center of Excellence (CoE)/Analytics Community

A center of excellence (CoE) is a team which provides leadership, best practices, research, support and/or training for a particular area of focus.

A CoE may also add value to an Organization through:

- Optimizing the organization by centralizing resources with high-demand and unique knowledge or skills and streamlining their contributions across a wide range of areas
- Improving ROI through the identification and development of reusable assets
- Reducing delivery times, development, and maintenance costs by increasing efficiencies
- Identifying and reducing duplication of effort across the enterprise

Interns

Acquiring seasonal interns is a long standing practice among many industries. For decades electric utilities have also used this recruiting method – relying on young interns to bring new ideas and skills into our industry. Not to mention, interns are vital in sustaining a stable workforce while battling natural attrition of tenured employees.

In years past, a student with a solid understanding of engineering topics and theories, with little exposure to computer or data science, would have been a viable candidate for a summer internship. While this might still be true for some of the industry’s core competencies, a student with a more rounded skill set will be needed in the world of data science. Luckily many of the engineering, business, and math curriculums have molded a solid foundation of computer programming & statistics into the minds of their students.

Identifying and retaining key talent is critical. The goal for both you and your intern is to have a productive and rewarding internship experience. Below is a list of suggestions that have led to success stories in the electric utility industry.

Identifying Talent

You will be hard pressed to find a “unicorn” during your hunt for interns. A unicorn being an individual who has both industry experience and a strong data science background. Within colleges and universities, students will most likely have focused on “projects” that are not necessarily industry specific.

When searching for the right candidate, listen to what the students have accomplished in their school projects. You will want a student that has coded in SQL, python, R, or at least C++. Listen to
how comfortable they are at sharing details. They should be able to speak clearly about more than just the end result their group delivered. Prod them about their specific role in the development of their project.

Along with programming knowledge, you will want a student that has a basic understanding of engineering principles related to your industry. A high level grasp of the science will suffice. Again, you should not expect a unicorn; but hopefully the candidate can conceptualize some fundamentals of your business.

Lastly, a strong math background is highly encouraged. Data science, forecasting, machine learning: these are all applications of statistics. Ask the students if they have heard or used some basic statistic principles in their projects at school.

Retaining Talent
In regards to training, if your group has the luxury of already having at least one solid data science engineer on staff, have the intern shadow this individual. Familiarity with the software is essential. If your group is lacking in data science expertise, online tutorials might be needed for your intern.

You will need to give the interns projects that are engaging. They need to feel that their contribution to your company is meaningful. In their minds, they are projecting this internship experience as a reflection of the company as a whole. Meaning, if they are given simplistic dull assignments, they will perceive the company as simplistic and dull as well.

Lastly, a fun work atmosphere is conducive to a harmonious experience. Be sure to include the intern in group outings, after work happy hours, and team building exercises. Engage the intern in water cooler conversations. Ralph Waldo Emerson once said “Nothing great was ever achieved without enthusiasm.”

Job Rotational Programs
Next Era Energy has stated that in the past “IT was expected to speak the language of Business; but now Business must learn to speak the language of IT.” Every department within any utility is now wholly reliant on IT support to maintain not only their computer hardware, but now their data as well. This melding of IT and Business needs to be galvanized in order to more effectively produce results. Next Era Energy is correct in thinking we need to learn each other’s languages at least at on conversational level.

Much like learning any new language, immersion into the culture is highly efficient. Rotational job programs can allow employees to experience other organizations within the company. During this experience, the employee will be exposed to the challenges unique to certain work groups. The knowledge the employees gain because of this rotational endeavor will create within them a better sense of cognitive business empathy. Moreover, in their future roles within the company, they will have the foresight to see the ripple effects of their business decisions and actions.
Partner Cross Functional Teams (IT & Business)

Communication is most effective when each party is speaking the same language. Looking back at the previous section on job rotations, a partnership between IT and business orgs can only be successful if they can communicate to each other the goals and challenges each face.

Setting Expectations from Educational Investments

Expectations of Data Analytics Resources

Developing Data Analytic skills is a time intensive endeavor. As stated above, Data Scientists, Engineers, and Analysts often have advanced degrees and have spent years honing their craft. To be successful in “growing your own talent”, a utility must be conscious of the time commitments required. Coincidentally, a prospective Data Analytics professional should be fully aware of the time commitments to become proficient. Success isn’t achieved through bare minimum levels of engagement.

Successful data analytics professionals never stop learning and improving their skills. They are constantly learning new techniques, researching new tools, are actively engaged in Data Science communities and professional organizations. Utilities should understand, respect, and encourage this behavior from their staff – it will pay off over time.

Measuring Results

When embarking on a Data Analytics training and development program, utilities should implement ways measure the program’s results. Identification of what methods worked, and those that did not (and for what reason) is important as your Education and Development program matures. Be sure to set clear expectations to your trainee’s as they begin the program. All parties should have a clear understanding of how success is defined.

Change Management

As with any adoption of new technology or process, change management is critical to success. Gaining acceptance of the use of data to drive decision making in some cases may be a daunting task. A strong change management strategy should be in place and include leveraging diversity, education/awareness training, communication strategies, and executive buy-in. In a sense, a roadmap should be developed to guide the organization through the new strategies and processes to ensure adoption. This section of the paper provides insight to each of these components of the roadmap for DA.

Gaining Executive Support

Strong leadership is one of the key drivers in an organizations ability to realize the power of analytics and move to a value-driving decision-making organization. Utilities successful in the use of Data Analytics have engaged management teams who understand the value DA can bring to the organization. A start to the process of gaining buy-in from the management team is to develop and roll out a vision of DA value. By focusing the vision on value, the DA results will have the greatest impact to the organization the desire and support for more will grow.
Leveraging Diversity
As discussed earlier in this paper, to develop accurate analytic models, multiple skillsets are required. Senior, experienced utility skills are needed to provide an accuracy check on the data and results. An experienced utility resource would be able to look at the data and results and know if it makes sense or not. They are also a great resource to develop ideas for use cases based on their work experiences in the company. Partnered with these experienced resources may be employees newer to the business, with very different skillsets including data mining, programming, statistical analysis, etc. Leveraging diversity means bringing together people of diverse backgrounds and experience level for a common cause.

Educating the Organization about Data Analytics (Create Awareness)
Sharing successes of other utilities, developing a library or catalog of use cases will aid in educating the organization and create awareness about Data Analytics. Consider integrating DA to an accepted business process/culture such as Six Sigma or Continuous Improvement. FPL/NextEra has had much success in tying formal training programs for DA to their Six Sigma culture. They have moved from IT personnel learns “the business” to Business learns IT. They use their Six Sigma culture and training process to tie together business knowledge, IT expertise and analytics skills.

Another way to educate the organization about DA is to conduct small group sessions with various organizational teams and share with them the vision of DA value and examples from other utilities. Then facilitate a working session in which the organization brainstorms ideas for possible use cases based on their new understanding of DA value. This risk in this option is that you may end up with
too many use cases resulting in an inability to meet the needs of the organization in a timely manner.

**Communication**

Communicating completed DA tools/models is part of “productionalizing” the results of the DA efforts. Productionalizing a model requires model validation, assessment, and maintenance. To aid in this effort, a Business Unit (BU) Owner of the model should be engaged from the start. The BU owner should provide input throughout the development of the DA model to ensure it meets requirements and provides valid, usable output. When communicating the completed DA model, the BU owner should be side by side with the DA team in describing the use for the model, the data sources, and how the model will be maintained. There may be various uses for the same use case and the BU owners should be able to speak to this clearly as the tool/model is rolled out.

**Building a Roadmap**

There are multiple facets to consider when developing a roadmap for Analytics including change management aspects. Start with need, not the solution. Engage with Business stakeholders early in the process to develop an effective communication plan. Understand where you currently are on the Data Analytics value curve and set a goal to where you want to be in upcoming years. Establish a shared, well-communicated vision prior to starting to work on specific use cases. Define business metrics which will be used to measure success and a means to prioritize business needs to ensure agreement on the best place to start. Once you have alignment with business strategies, the leadership team, and the overall goals of analytics, the next step is to focus on the skillsets required.

An enablement section of your roadmap should describe how you will enable the organization to grow analytics capabilities to move along the Data Analytics Value curve. What types of training, communication, skillset development across the enterprise should be put in place to ensure adoption and expansion of data analytics in multiple business units? How will you structure the organization to enable a focus on analytics? Will you enable self-service analytics? Consider specific change management programs which need to be in place to democratize analytics and analytic capabilities.

The last section of the roadmap should layout a means to provide optimization and scaling of analytics functions. This section analyzes the current IT architecture for data & analytics tools and proposes a strategy which supports future growth as analytics spread throughout the company. This section may present a multi-year growth model for IT architecture to add features (e.g., cloud technology, data warehouses, governance, etc.) over the course of a specified period of time.

The challenge is to keep up with the pace of change: technology changes, organizational changes, personnel changes, and the utility’s natural desire to build upon the successes realized through Analytics.
References


https://www.kdnuggets.com/2018/05/simplilearn-9-must-have-skills-data-scientist.html