A New Jersey Big Data Workforce Roadmap: An Examination of the Challenges and Opportunities for New Jersey's Workforce to Successfully Compete in the Data-Driven Economy

Submitted by the New Jersey Big Data Alliance

Peggy Brennan-Tonetta
Rutgers University

Deborah Silver
Rutgers University

Jessica Paolini
Rutgers University

Rashmi Jain
Montclair State University

Ethne Swartz
Montclair State University

Marina Johnson
Montclair State University

Stanislav Mamonov
Montclair State University

Chelsey Hill
Montclair State University

Introduction

The research team representing the New Jersey Big Data Alliance examined advanced computing, big data and analytics technologies, and skill sets that will affect the future of work in New Jersey, and the type and extent of the impact of these changes on key industry clusters in the state. The result is an evidence-based roadmap for New Jersey to prepare workers and industry to successfully compete in the data-driven economy. For purposes of this report, big data refers to extremely large data sets, as well as the technologies, people, and processes that allow for its collection, storage, and analysis. Multitudes of data are generated every second from sources all around us. Computational analysis of these data enables companies to extract new insights (such as revealing patterns and trends) and create new forms of value. It is critical for companies and workers to adjust to, and prepare for, the changing technological environment driven by big data that is now key to economic revitalization. In fact, big data has become so pervasive and the opportunities it presents so transformative, that the ability to analyze it and glean valuable information is now viewed as essential for competitive growth.

To fully maximize the benefits of the new data-driven economy, companies need access to high-end computing technologies and workers who are computer savvy and proficient in analytics. Industry, across all sectors, has a strong demand for data science talent and struggles to fill open positions. According to Burning Glass Technologies, data science and analytics (DSA) jobs remain open 13% longer than non-DSA jobs (Sigelman, 2018).

The team utilized existing research and conducted primary research to identify workforce trends, gaps, and opportunities, and to provide recommendations that will enable workers and industry to utilize new technologies to analyze and glean information from big data. Specifically, the research team provided a description of current and future advances in technology as it

relates to big data, artificial intelligence (AI), and advanced computing, examined how these advances are changing the landscape of New Jersey's economy, and examined the impact on the workforce, with a focus on how they are reshaping key industry clusters in the state[1] (e.g., health care, logistics, food, financial services, clean energy and advanced manufacturing). A survey and analysis was also conducted of existing data analytics and computing programs in the state. The report concludes with recommendations for effective strategies tailored to various actors such as universities, state agencies, and employers in educating, training, and equipping New Jersey's workforce with the skill sets necessary to succeed in a data-driven environment.

Overview of Technologies Affecting Workforce Today and in the Future

The Third Industrial Revolution, which started in the 1980s, ushered in digitization that facilitated automation and efficiency. The Fourth Industrial Revolution, which builds on the Digital Revolution and is now taking shape, is an age of the convergence of physical, digital, biological, and chemical phenomena (World Economic Forum, 2018). The result will be a deluge of new forms of innovation, driven by disruptive technologies, connectedness (Internet of Things - IoT), and big data. These changes are creating a demand for a workforce with the skills to understand new ways of approaching the innovation process and the technologies required to realize its full potential. State governments and industry sectors that understand the possibilities that result from this convergence early on, will establish a competitive advantage and be at the forefront of Industry 4.0 economic growth.

AI and machine learning are gateway technologies to the Fourth Industrial Revolution (Lee, Davari, Singh, & Pandhare, 2018). Although a universally accepted definition of AI remains elusive, a study on *Artificial Intelligence and Life in 2030* suggests that AI generally refers to the multi-disciplinary efforts devoted to making machines intelligent, i.e. making

machines function appropriately and with foresight (Stone et al., 2016). The McKinsey Global Institute (2017) regards AI as a "family of technologies" that comprise advanced machine learning, computer vision, natural language processing, virtual assistants, and robotic process automation, among others.

AI applications are supported by decades of research on machine learning, computer vision, and natural language processing, and they are making a significant impact on almost every aspect of human activity by improving health, safety, and productivity. The global economic impact of AI is expected to reach $15.7 trillion by 2030 (Rao, Verweij, & Cameron, 2017) and AI capabilities development has been recognized as a national priority in the United States (The White House Office of Science and Technology Policy, 2018).

Although much of the excitement about the AI potential is associated with expected future developments, AI is already making an impact. The level of adoption of AI is in part dependent on the adoption rate of digital practices across different industries. Focusing on the AI impact across specific verticals, an analysis conducted at PricewaterhouseCoopers indicates that health care, transportation, and finance are the top three industries that are expected to benefit from AI (Rao et al., 2017). In terms of health care, AI is expected to have a significant impact in improving patient outcomes and reducing costs (MITRE, 2017). For example, a recent announcement from Google that an AI system developed by the company is significantly better than human dermatologists at detecting cases of melanoma is indicative of the kind of improvements that are expected (Google, 2018).

While there is a general agreement on the transformative effect of AI applications, industry research provides varied perspectives on the effects of AI on the labor market. A recent survey of executives across different industries conducted by the McKinsey Global Institute

(2018) suggests that most industries do not expect significant changes in employment as a consequence of the AI adoption in the near term. However, another recent survey conducted by Deloitte suggests that the leading companies across different industries are making substantial investments in AI and expect a significant reduction in personnel as a consequence of expanding AI adoption (Loucks, Davenport, & Schatsky, 2018).

Despite the disagreement on workforce impact as a result of AI, there is consensus across the surveys that the companies making investments in AI are facing significant challenges in recruiting employees with the required AI skills. In a recent report by Indeed, a national job search firm, data indicates the shortage is getting worse: "While more job seekers are interested in data-science jobs, the number of job postings from employers has been rising faster than the number of interested applicants" (Perry, 2019), with the New York region experiencing the biggest gap in talent. LinkedIn calculates that, in August 2018, employers nationwide were seeking 151,717 more data scientists than exist in the United States (Perry, 2018).

With regard to specific AI skills that are sought by companies, Shoham et al. (2018) cite the top two skill sets as machine learning and deep learning, followed by natural language processing, robotics, computer vision, and speech recognition. While machine learning appeared to be the most common skill cited in job listings, deep learning experienced a 35x growth from 2015-2017.

AI talent shortage has been recognized as a key impediment in leveraging AI for economic growth at the national level. The White House report on AI suggests that sector-specific focus in developing AI talent would be advantageous (The White House Office of Science and Technology Policy, 2018). This preliminary recommendation is in agreement with strategies that are being developed by other countries. A study assessing the potential for the

United Kingdom (UK) to capitalize on the opportunities created by AI suggests that the following elements provide an advantage: dynamic research community, vigorous start-up ecosystem, and a constellation of legal, financial, and linguistic strengths located in close proximity (Select Committee on Artificial Intelligence, 2018). The UK report also includes a number of specific tactical recommendations that could be adopted in the New Jersey context: industry sponsored Ph.D. positions that focus AI research on industry topics, retraining programs for workers displaced by AI, targeted adoption of AI in the public sector (e.g. health care), prioritized AI adoption at the state level, leveraging AI for government services, and making data sets available for workforce training.

AI is also regarded as having a significant impact on cybersecurity (Sikos, 2018), both positive and negative. From the negative perspective, the rise of AI systems that are "super-intelligent," coupled with cyber attacks sanctioned by government actors, could result in catastrophic failures of key infrastructures. The frequency and sophistication of cyber attacks is expected to intensify as hackers deploy machine learning software and AI. IBM researchers warn that AI can be used to develop new malware, making it harder to detect and more effective at reaching its target (Krause, 2018).

From a positive perspective, AI can also be used to respond to and prevent cyber threats. For example, AI can speed up "incident responses once malicious software is detected on computer networks" and "help thwart email-delivered ransomware or swarming botnets" (Krause, 2018). Google relies on AI to protect its data centers and Microsoft's anti-malware service, Windows Defender, is based on machine learning algorithms. To reduce the potential for AI to be executed in a cyber attack, it is "important for security companies — and their customers — to monitor and minimize the risks associated with algorithmic models" (Giles,

2018). According to Martin Giles (2018), San Francisco bureau chief of *MIT Technology Review*, this requires people with "deep expertise in cybersecurity and in data science". People with this combination of expertise "are still as rare as a cool day in a Las Vegas summer".

P&S Market Research expects the cybersecurity AI market to grow by 36% each year from 2017 through 2023 as the computer security industry turns to AI to more efficiently detect and react to threats (Giles, 2018). AI is also viewed as a solution for addressing the "massive shortage of skilled cyber workers" (Giles, 2018). By the end of 2016, more than 22,000 cybersecurity workers were employed in New Jersey, but the "demand for cybersecurity talent outstrips the supply of available workers" (New Jersey Business Magazine, 2017). Incorporating AI in security applications may lower the long-range forecast for cybersecurity talent, while increasing the number of hybridized roles that require workers to have skills and experience in cybersecurity *and* data science.

Impacts of Technological Advancements on Key Industry Clusters

The following discussion of the impacts of technological advancements on New Jersey industry is organized around priority clusters in the state as identified in *The State of Innovation: Building a Stronger and Fairer Economy in New Jersey* (New Jersey Business Magazine, 2017). These clusters include: Life Sciences; Advanced Manufacturing; Transportation, Logistics, and Distribution; Finance; Food and Agriculture; and Clean Energy.

Life Sciences

In New Jersey, the manifestations of the Fourth Industrial Revolution are becoming evident in the transformation of one of the state's core industries — life sciences. Factors leading to change include new opportunities to identify and diagnose diseases made possible by the

convergence of technologies, the disruption driven by start-ups, and by technology companies such as Amazon, Microsoft, and Google entering the space.

Advances in big data tools and analysis offer new opportunities for accelerating drug innovation, improving determination of clinical trial outcomes more quickly, identifying unmet medical needs, gaining access to new markets, and defending drug pricing and improving margins (Oracle, 2015). According to GlobalData, pharmaceutical executives identify big data analytics and AI (in order) as the technology trends that will have the greatest impact on their business in 2019 (Donovan, 2017).

The Food and Drug Administration (FDA) is actively developing new protocols, data sets, and programs to leverage advanced computing technologies for human health. FDA Commissioner Scott Gottlieb, M.D., recently shared that the FDA is writing a regulatory framework for the use of AI in medicine (Kent, 2018). The agency also aims to leverage "big data to accelerate clinical trial efficiency, [and] medical product development" (Kent, 2018). This change is already happening in New Jersey. BioNJ (2018) reports that there is substantial demand among life sciences employers for data scientists and engineers to support clinical trial optimization and other activities. New Jersey can further solidify its position as a top location for clinical trial activity and accelerate economic growth by encouraging the adoption of big data technologies in the clinical trial life cycle and training/up-skilling workers for in-demand data science and engineering roles.

New combinations of biotechnology, bioinformatics, and nanotechnology are enabling life sciences companies to leverage powerful computing power and big data analysis to understand disease states and offer new solutions that are purely digital, bypassing conventional drug treatments. For example, GlaxoSmithKline made significant investments in 23andMe,

spurred by a desire to participate in a new era of drug discovery — genomics. The two companies will collaborate to utilize 23andMe genetics data to help with drug discovery, clinical trial management, and recruitment, while also using machine learning to guide the selection of drug and disease candidates (23andMe, 2018). The field of genomics research has seen unprecedented growth and innovation due to the application of big data analytics, and unlikely players are now entering the field. Increasingly, Google and Amazon have begun to make investments in these areas as well (Monegain, 2016).

The medical devices and equipment subsector is increasingly relying on advanced computing technologies, data analytics, and predictive learning to drive innovation (Kent, 2018). To increase market share by developing and leveraging new medical technologies, the New Jersey medical devices subsector will need access to digital talent. Deloitte notes that the medical technology workforce "is evolving from one that historically focused on electrical and mechanical engineering and product development, to a new and flexible class of workers who possess skills in digital, advanced data analytics, and machine learning" and recommends that to remain competitive, companies will need workers with experience in app and algorithm development and "multidisciplinary talent from creative and scientific backgrounds (e.g., advertising, statistics)" (Arboleda, 2018). Key stakeholders in the medical technology subsector are concerned that the widening skills gap will hinder market growth (Arboleda, 2018).

Advanced Manufacturing

Advanced manufacturing is a significant sector in New Jersey, employing highly educated workers in a variety of industry verticals. Eighty-one percent of all manufacturing in New Jersey is now classified as advanced (Arboleda, 2018). These employers seek high-tech manufacturing skills such as computer-based modeling, 3D printing capabilities, and analytical

skills that require modern big data approaches to analyze and model complex data, including R, Python, and data visualization tools in data management roles. Manufacturing is being reinvented with smart devices, automation, and big data analytics, and as a consequence, new jobs are being created. New growth in manufacturing employment is being created in knowledge-intensive roles focused on the creation, implementation, optimization, and continuous improvement of manufacturing systems. Thus, the key to growing manufacturing is not to "bring back" the jobs that left 20 years ago, because those low-skill, labor-intensive jobs are disappearing everywhere. The key is to grow manufacturing anew, by utilizing advances in technology to create systems that are more productive, more effective, and more agile, and to have a workforce trained in the skills necessary to fully utilize and apply technological advancements.

Finance

New Jersey has become an attractive location for companies in the financial technology industry that utilize technology innovations in next-generation banking, credit, and investment services. Mobile Internet and cloud technology are already affecting the way the finance industry works. AI and data mining are still in their early stages of use, but the pace of adoption is rapidly increasing. Concurrently, the growth of Blockchain technology will improve transaction security when dealing with sensitive data. Enabling these advancements, New Jersey provides one of the most extensive fiber optic networks in the world, which helps support the financial industry's high-speed communications needs. The state's IT infrastructure has resulted in the second largest concentration of data centers in the nation. With more than 100 in northern New Jersey alone, these servers power the lifeblood of high-speed electronic trading. Financial exchanges such as the New York Stock Exchange and its competitor Nasdaq OMX now process the vast majority of

their trading information in server farms in New Jersey. Data analytics as a distinct competitive advantage will continue to grow and fuel New Jersey's data center market, resulting in high demand for workers with this skill set. Currently, deep learning is the AI area in most demand, encompassing all the neural network-related techniques, and "people who have expertise in knowing how to develop those new techniques, these topologies, or how to implement them in the most efficient manner in software and hardware obviously have high value" (Perry, 2019),

Transportation, Logistics, and Distribution

New Jersey is home to one of the world's fastest-growing hubs for transportation, logistics and distribution (TLD), with a projected growth rate of 42%. According to industry and occupational projections, the top 10 TLD industry cluster occupations in New Jersey (by employment) are expected to see an employment increase of over 9% from 2016 to 2026 for a total of almost 27,000 added jobs (New Jersey Department of Labor and Workforce Development, 2018). However, the sector is dealing with major workforce issues such as an aging worker population and need for retraining, up-skilling, and education.

In addition, the current TLD workforce in New Jersey is comprised of over 50% of workers 45 years or older (25% are over 55 years old), and a similar percentage having only a high school diploma or below educational level (New Jersey Department of Labor and Workforce Development, 2018). This, coupled with continuing growth projections in employment, calls for a "reimagination" of the TLD sector. TLD companies are leveraging big data and analytical tools to improve levels of service and operational efficiency, and a skilled workforce is needed to utilize these technologies. Examples include business intelligence tools to predict route needs related to transport vehicles, crews, and facilities; proactively manage vehicle maintenance; cost-effective inventory management; and improved understanding of and

responses to safety issues. Thus, the nature of employment will undergo tremendous changes, redefining the growth potential in this sector. Employment is expected to increase, driven by the growth of e-commerce and international trade, as well as a significant transformation driven by robotic process automation, IoT, and application of AI, fueling efficiencies and growth. UPS, the world's largest package delivery company, is an outstanding example of this. Today, it is considered a technology company in many ways due to its innovation in applying big data analytics and IT solutions.

Food and Agriculture

Food companies of all sizes and types are embracing digital transformation across their operations. The adoption of automation, data integration, analytics, simulation, and energy sourcing can help companies transform existing plants into smart manufacturing facilities. In addition, changes in the marketplace are driving tremendous innovation, resulting in new types of functional/health conferring ingredients, new forms of packaging, new methods of food/agricultural production and distribution, and a greater diversity of products than ever before. In terms of the agricultural sector, precision agriculture technology such as sensors in fields that provide insight into fertilizer requirements and disease pressure give farmers data-driven information never before available, while radio-frequency identification and global positioning system technologies allow for the tracking of products as they move through the supply chain, and analysis of large volumes of consumer data make it easier to forecast demand and reduce food waste (Irani et al., 2018). The next generation of farmers will need to have more than the traditional expertise in plant and soil science, but also knowledge of geographic information systems and IoT to connect agricultural operations, and automation of production practices for example, which will require skill sets in programming, analytics, and computing.

The key to fully realizing the value of these opportunities is to create interconnected operations and supply chains through the application and utilization of a broad array of technologies and analytics capabilities, thereby catalyzing a new era of smart food — from farm to fork.

Clean Energy

New Jersey has a goal to develop 600 megawatts of energy storage systems by 2021, and 2,000 megawatts by 2030. Several battery storage projects have already broken ground in Ocean, Atlantic, and Warren counties. Big data and machine learning facilitate "intelligent" energy storage management through analysis of weather patterns and second-by-second energy consumption data to determine the ideal times for charging (off-peak, when rates are lowest) and discharging (during peak demand) energy storage batteries. In a recent presentation, Matthew Sachs, Chief Operating Officer at Peak Power, stated that machine learning-enabled forecasting "maximizes the use of intermittent renewable power" like wind and solar (Sachs, 2018). Intelligent energy storage management is a crucial component of New Jersey's clean energy economy. Workers with the ability to analyze energy data streams will be key to realizing the benefits that energy storage offers.

Skill Sets Needed for the Future Workforce

This section explores how business needs and technology evolution drive the need for talent in the dynamic New Jersey/New York region. Companies are shifting from generating data to relying on data to make decisions, and creating business value in the process (Berinato, 2019). Berinato (2019, p. 128) suggests that analytics projects add value when a team "… asks smart questions, wrangle relevant data and uncover insights. Second, it must figure out — and communicate — what those insights mean for the business. The ability to do both is extremely rare…" and requires a variety of capabilities, including project management, data analysis, data

wrangling, design, and storytelling. Additionally, subject-matter expertise is critical to help the team focus on business outcomes.

Big data jobs exist along a dynamic "pipeline" of capabilities. Burning Glass Technologies, an analytics software company that provides real-time data on job growth, skills in demand, and labor market trends, suggests that these capabilities span entry-level, required skills, defining skills, and distinguishing skills (which are the most cutting edge). To understand this continuum, a search was conducted by the team using the Burning Glass Labor Insights platform[2] (LaborInsight, 2018) that looked at job postings data.

At the top of the jobs funnel, there were approximately one million job postings (about 40,000 in New Jersey) that Burning Glass classified as big data jobs for the period January to December 2018. This included many traditional jobs that are now transitioning to include big data skills, such as all data and financial analysts, business consultants, database developers, and software engineers. The skills needed today for these jobs are mostly Excel, Structured Query Language (SQL), budgeting, project management, and general business knowhow, but this will change. In the future, many of these jobs will require the cutting-edge or distinguishing skills that high-tech data science jobs are increasingly requiring.

What are these cutting-edge job opportunities and what skills are needed? To focus on just this segment, the Labor Insight search was set to only those job postings where the title reflects these new 4.0 careers (examples of these titles include Data Engineer, Data Scientist, Business Intelligence Developer, Machine Learning Engineer, Analytics Consultant).[3] The data showed approximately 182,000 jobs were advertised for positions with titles that included analytics and data science in the United States over the past 12 months. The New Jersey/New York MSA had the highest number of postings (20,000). At a state level, New Jersey had about

7,000 job postings.[4] Skills requested in the job advertisements included SQL, but also Python, machine learning, Hadoop, business intelligence, among others. The top skills are listed in Table 1.

[Insert Table 1 About Here]

In addition to these specialized data skills, general business skills were also required in these job postings. These skills include communication, teamwork, research, and problem solving, among others. A complete list is provided in Table 2. All of these skills, both the specialized technical skills and general business skills, are the same skills identified by LinkedIn as the most in-demand skills for 2018 (Bowley, 2018).

[Insert Table 2 About Here]

As we think about preparing a workforce for the Fourth Industrial Revolution, hybrid skill sets are at the forefront. This combination has been reported on recently in the Wall Street Journal (Weber, 2019), "The Hybrid Skills that Tomorrow's Jobs Will Require," which summarized the Burning Glass (2019) report on "The Hybrid Job Economy." To quote from the report, "Once considered highly technical jobs, now these jobs require writing skills, problem-solving skills, creative and research skills, and skills in teamwork and collaboration. So just like the marketing manager who is now an analyst, the software engineer or data scientist is now a business person, designer, and team worker." The report describes major skills that all workers need in order to adapt to the changing workforce. These include: a comfort level with digital tools and technology, an ability to handle data and analytics, an understanding of business and management, and a design/creative outlook. The hybrid skill sets, combining technical expertise with business and communication skills is especially important in data science and analytics careers.

In order to obtain an overview of the type of employers that are seeking to fill big data jobs, clusters were created in the Labor Insight database that correspond to the key New Jersey clusters identified earlier in this report. Table 3 shows the Burning Glass Labor Insight analysis results nationally and for the New Jersey/New York region. It is worth noting that the Harvard Business School cluster study for this region shows the dominance of the business services cluster that corresponds to the professional, scientific, and technical services cluster in Table 3.[5]

[Insert Table 3 About Here]

The skills required in these individual sectors were also examined to see if there were any significant differences. The professional, scientific, and technical services industry most follows the skill listing seen in Table 1. In the finance category, not surprisingly, Statistical Analysis System, economics, and risk management are higher up on the skill requested list. In the information sector, more specialized skills are required such as Scala, C++, and software development being requested, as these jobs require more programming and are traditionally hiring in the software space. Doing a deeper dive in terms of the pharma and chemical industry, data science job postings in these areas also request market research, biotechnology, and simulation expertise. Finally, in the advanced manufacturing space are robotics, C++, simulation, and computer vision, all reflecting more engineering focus. See Silver, Bernis, and Weston, 2018 for an analysis of big data skills by sector for entry-level jobs.

Burning Glass data also show that the demand for IT and cybersecurity jobs is high in New Jersey, second only to the Washington, D.C. area (Cleary, 2017). The skills that are in very high demand in New Jersey are in risk management, audit planning, risk assessment, certified information systems auditing, accounting, and internal auditing as these relate to cybersecurity. All these were occupations with high Location Quotient scores (ranging from 1.6 to 2) for the

New Jersey/New York area. Information security is a position that has a lower Location Quotient but is the most common job title, comprising of 39% of all jobs in information and cybersecurity. Of note is that there is clearly a need for raising awareness about the importance of certified information systems auditing certification, or similar certifications.[6]

What are the fastest growing skills across these industries? It is expected that employers will increase their need for predictive analysis, machine learning, and data visualization (which is why knowledge of Tableau or an equivalent tool is essential). Burning Glass predicts an increase in demand for more nuanced skills such as NoSQL for data handling and R for data analysis (Silver, Bernis, & Weston, 2018). Other software skills anticipated to increase in desirability include Sqoop, used to import data from relational databases, and Apache Hive/Hadoop for data summarization, query, and analysis. Data analytics skills required to extract information and intelligence essential for understanding the global marketplace will continue to grow in importance. Unstructured and real-time data will provide companies with a competitive advantage and individualize the client experience.

CyberSeek, a free career resource tool created by CompTIA and Burning Glass Technologies, identifies career pathways for cybersecurity professionals. The top skills requested for cybersecurity roles were analyzed by the team to identify crossover skills with application in data science. The team reviewed 10 jobs classified by CyberSeek as "core roles," including three at the advanced career level (cybersecurity manager/administrator, cybersecurity engineer, and cybersecurity architect); three at the mid-career level (cybersecurity analyst, cybersecurity consultant, and penetration and vulnerability tester); and four at the entry level (cybersecurity specialist/technician, cybercrime analyst/investigator, incident analyst/responder, and IT auditor).

The top requested skills for cybersecurity jobs were compared to Burning Glass data on in-demand data science skills. Most of the cybersecurity roles (7 of 10) require skills in project management, which is requested in 13.6% of open analytics and data science jobs. There are also crossover skills in programming languages, including Python and Java. Cybersecurity engineers, cybersecurity consultants, and cybercrime analysts/investigators are expected to have competency in Python, a skill listed in nearly 40% of all data science job postings. Penetration and vulnerability testers (also commonly referred to as penetration tester, application security architect, application security analyst, and security analyst III) have the most crossover skills: Python, Java, Linux, and project management.

Additionally, networking jobs, which often serve as feeder roles to careers in cybersecurity, typically require knowledge of Python and Java. This is worth noting, because the New Jersey Department of Labor and Workforce Development projects job losses for network and computer systems administrators and computer network architects (employment change of -0.2% and -2.0%, respectively) (New Jersey Department of Labor and Workforce Development, 2019). Workers in networking jobs, or students preparing for networking roles, should consider broadening their focus to incorporate knowledge of cybersecurity and data science. This combination of skills will make them eligible for a range of in-demand roles, in cybersecurity, data science, and cybersecurity AI.

In November 2018, the New Jersey Big Data Alliance hosted a Data Science Workforce Forum. More than 30 individuals participated in the event, representing industry, academia, and government. Employers were asked to complete a survey about their hiring needs, workforce training programs, and technology trends. A majority of the employers that completed the survey identified machine learning as an emerging technology trend expected to have a significant

impact on their business or industry within a three to five year or five+ year timeframe. Half of the employers indicated that they already recruit or plan to recruit candidates for hybridized roles (subject-matter expertise or industry-specific experience + technology skills). Several companies emphasized the need for all workers to be "digitally savvy" with one employer writing that "…being able to employ analytics skills in any job offers professionals a leg up in the ultra-competitive market we exist in today. Comparing apples to apples, the candidate that has basic to advanced competencies in digital literacy has the edge when competing for employment opportunities."

Analysis of Programs and Curriculum for Preparing a Future Workforce

As previously discussed, almost every job from new hires to C-suite will require some level of data analytics skills. To meet this demand, many colleges in New Jersey have developed new programs incorporating big data and advanced computing concepts into undergraduate and graduate curriculum, and started offering Massive Open Online Courses and online degrees in the field of data analytics to make big data education more accessible (Cegielski & Jones-Farmer, 2016). Primarily, two interdisciplinary fields — data science and business analytics — have recently emerged to address the talent shortage in the job market in the United States, as well as in New Jersey.

Data science and business analytics degrees are primarily offered through computer science departments, as well as business schools. Many colleges are partnering with industry to improve their courses and provide students with research experience. For example, Google will open an AI lab in Princeton and will collaborate with university researchers. Although some schools offer interdisciplinary programs and find ways to incorporate computer science, business, information systems, statistics, etc., many universities offer data science and business

analytics in disciplinary silos. This is an issue that needs to be addressed because data science is a multifaceted discipline.

Although fully analyzing big data to extract information necessitates advanced computing skills, a study regarding the requirements of analytics programs find that existing analytics curricula do not always require advanced programming courses (Aasheim, Rutner, Williams, & Gardiner, 2014). According to a study conducted by the National Academy of Sciences, Engineering, and Medicine (2018)**,** some of the essential advanced computer skills required to handle big data are: accessing databases to query data, ingesting data into databases, gathering data from other sources using web scraping, parsing, and tokenizing texts in the environment where big data are stored (e.g., the cloud). Primary concepts that students should know regarding computer programming are logic, data structures and algorithms, and abstractions. It is also important for students to learn new technologies, programming languages, and software applications that help store, process, and analyze big data such as Python, Hadoop, and Amazon Web Services (Cegielski & Jones-Farmer, 2016).

Educators and higher institutions should also consider that every student, regardless of their major, should have some exposure to data science. The data science programs offered through computer science departments in the tri-state area (New Jersey, New York, and Pennsylvania) offer programming courses. However, few undergraduate and graduate programs in business analytics offer programming courses. It is important that institutions and departments in the state emphasize the importance of programming skills and provide opportunities for students to learn them. **There should also be a vision for K-12 data science education in New Jersey.**

New Jersey institutions educating big data analytics practitioners should expose students not only to structured data, but also to unstructured data, including text, images, and videos. Students also need to acquire practical experience to query data from both relational and non-relational databases. This, of course, requires use of big data sets in the classrooms. The New Jersey Office of Information Technology Open Data Center provides a great resource for faculty looking for data sets to use in their classrooms.

To describe and analyze big data, visualization and modeling skills are essential. Big data analytics practitioners who create reports and dashboards need data transformation, exploratory data analysis, and descriptive modeling skills. Analytics practitioners who are responsible for building both predictive and prescriptive models need to know machine learning algorithms, as well as linear and non-linear optimization techniques. Students should experience the entire big data analytics life cycle: data acquisition using SQL, data transformation, feature selection, model building and assessment, and putting the best model into production and use.

Communication and presentation skills are also essential because big data analytics professionals extract information from large data sets and present it to decision-makers. The ability to present the high-level insights of an analytics project in a clear, concise manner is crucial. For this reason, big data analytics practitioners should be able to simplify the results of their analyses and convey only the most important points and insights to senior-level managers.

The nature of data analytics is multi-disciplinary and team members with diverse backgrounds exchange ideas and information. Therefore, data analytics education needs to also incorporate modules in which students are required to communicate with other people within or outside of their team to ensure success in the workforce. Class projects that utilize

multidisciplinary teamwork can be beneficial for students to improve their presentation skills and obtain business acumen, as well as domain knowledge.

The multidisciplinary and interdisciplinary nature of data science is well suited for incorporating data science education across disciplines. As the importance of data science grows, it is becoming increasingly necessary for students to acquire these skills, regardless of their chosen field of study. **For this reason, the implementation of data science/analytics as a general education requirement in New Jersey colleges and universities is an indisputable necessity for the future workforce.**

The study team collected and examined a sample of data science and business analytics course offerings from 26 four-year colleges and universities in the tri-state area. A summary of this information can be found in Table 4. Based on an analysis of programs in New Jersey, there are two major gaps in data science education that must be addressed for tomorrow's workforce to meet the evolving data science needs of industry in New Jersey: 1) cross-program course offerings, and 2) bridging the divide between theory and practice.

[Insert Table 4 About Here]

The majority of big data analytics programs provide courses in the focus areas of the aforementioned components of data science and analytics; however, what is often lacking in these programs is the exposure of students to real data sets and business problems. A survey conducted by Wixom et al. (2014) found that industry is not satisfied with graduates' practical experience, due in large part to a lack of focus on real-world applications of analytics using real data sets. Additionally, according to Tang and Sae-Lim (2016) , while the term "experience" is one of the most frequently used terms in the descriptions of the 20 data science programs in their sample, many of these programs lack a capstone course. Capstone courses are one way in which

data science programs offer students hands-on, real-world learning experiences. In addition, many programs also do not require internships or co-ops to graduate, further limiting student exposure to real-world analytics. Figure 1 represents a SWOT (strengths, weaknesses, opportunities, and threats) analysis diagram, which summarizes the aforementioned information regarding data science and business analytics higher education program offerings in the tri-state region.

[Insert Figure 1 About Here]

Co-curricular/extracurricular offerings can help to bridge the gaps that exist in data science education. Seminars can focus on data science skills and techniques that are not explicitly taught in the classroom. Bringing industry professionals into the classroom or face-to-face with students can be an invaluable tool to offer new insights and perspectives to students. Whether inside or outside of the classroom, opportunities for students to network, communicate, and learn from one another is crucial to their development as data scientists. Team-based workplaces are prevalent and future data science practitioners must be prepared. Extracurricular programs, such as student-run and faculty-advised data science and big data clubs, hack-a-thons, etc. offer additional collaboration and learning opportunities for students.

Conclusion and Recommendations

Talent and human capital are now, and will continue to be, the critical factors of production (World Economic Forum, 2018) in the data-driven economy. Technology and lifelong training to maintain the appropriate skills become pre-conditions for participating in this new landscape. It is predicted that the use of big data and AI will transform economies and deliver a new wave of productive growth (Catlin, Scanlan, & Willmott, 2015). However, most companies have not realized the full potential of these technological advances due to a number of

barriers, including talent shortages. By educating, training, and facilitating access to individuals with advanced computing and analytics skill sets, New Jersey can provide a competitive advantage for it employers. A "Big Data Workforce Roadmap" that describes the current and future skills needed as a result of technological advancements big data and AI, and roles of actors in the workforce development space, is provided in Table 5. Additional specific recommendations are as follows:

[Insert Table 5 About Here]

Data Science Education and Training

- Create a vision for K-12 data science education in New Jersey.
- Require all high school students to take an introductory class in programming logic/coding. Some private schools in New Jersey are introducing programming languages like Python through summer academies or through after-school clubs. New Jersey Big Data Alliance institutions can work with high school teachers to support the development and teaching of suitable curriculum.
- All New Jersey two- and four-year colleges/universities should require data science and/or analytics as a general education requirement for graduation.
- Track enrollment in AI courses in colleges and create a competitive statewide summer academy for the most promising students to participate in research fellowships and mentorships with key professors. This currently happens with the Governor Schools for high school students; an equivalent is needed for college students.
- Evaluate in-demand certifications and redirect students toward in-demand and emerging qualifications. For example, Cyberseek.org projects that the New Jersey cybersecurity workforce will experience a shortage in certifications such as certified information systems

auditor, certified information security manager, and certified information systems security professional, and an over-supply of lower-level certifications in computer hardware management qualifications.

- Develop big data project management training modules for mid-management-level workers who may not need an in-depth knowledge of data science, but who are responsible for identifying the scope of big data projects, resources, and risks associated with such projects, designing, and implementing them. Often these managers oversee data analytics teams and need to have at least some basic understanding of data science terminology, methodologies, and the various applications for gleaning data insights. As an example, the New Jersey Big Data Alliance will be hosting a workshop on managing data projects for senior executives.
- Develop programs to transition highly educated researchers, scientists, and business professionals dislocated when companies move out of state. Many of these mid-career to advanced career-level professionals have the necessary capabilities and professional qualifications to acquire new big data skills.
- Capitalize on big data insights to improve education and training. As an example, Georgia State is using big data and predictive analytics to inform and supplement more traditional services such as student advising. School administrators credit their big data technologies with increasing graduation rates, as well as identifying factors that can predict student success in specific programs (for example, a grade in a student's first math class was more predictive of success in the nursing program than grades in physiology or anatomy).

Industry/Academia Collaboration

- Facilitate stronger connections between data science talent and New Jersey employers through advisory boards, data science-focused career fairs, a career pathways database,

industry-led student workshops and mentoring programs, and mock interviews. These provide employers the opportunity to assess the hard/soft skills of students, while providing students with industry contacts, mentors, business knowledge, and interviewing experience.

- Provide hands-on experience opportunities for students to learn about the application of data science across a range of industries and job functions. Create a Data Science Apprenticeship Program that allows students to rotate among diverse (i.e., different industries) companies.
- Give employers time and a forum for sharing knowledge with students. This could be achieved through hands-on student projects, mentoring, and/or student interviews with advanced career-stage workers to capture employee knowledge.

Policy Frameworks

- Universities in Canada have been encouraged by government to teach AI courses, and establish fellowships and research institutes that collaborate with industry. There is a portal that advertises AI start-ups and talent (http://www.canada.ai/directory) that acts as a magnet for drawing yet more talent. New Jersey economic development organizations (EDOs), such as Choose New Jersey, should establish a directory of AI start-ups to use as a marketing tool for attracting new companies, investors, and talent to the state.
- New Jersey EDOs can follow the lead of EDOs from Texas to North Carolina that are focusing less on company attraction and retention and more on talent recruitment and retention, based on the recognition that **access to talent is a top location decision criteria for corporations.**
    - State provides funding to recruit Nobel Laureates to Texas higher education institutions.

    - Marketing campaigns in various states provide information about local employers, job opportunities, Quality of Life indicators, and encourage networking among young professionals. Example: North Carolina - Research in the Triangle, Smarter from Any Angle.

- Countries with leading AI sectors have clear AI or Fourth Industrial Revolution strategies for innovation (Denmark and Germany are good examples). New Jersey should develop an AI/Deep Learning education and industry development strategy that involves relevant state departments, higher education representatives, and key employer organizations.
- A recent report by the UK includes a number of specific tactical recommendations that could be adopted in the New Jersey context: industry-sponsored Ph.D. positions that focus AI research on industry topics, retraining programs for workers displaced by AI, leveraging AI for government services, and making data sets available for workforce training (Select Committee on Artificial Intelligence, 2018).

Table 1. Top Big Data Titled Jobs: Skills In Demand (January to December 2018)

Skills	Job Postings
SQL	44.2%
Python	38.9%
Data Science	32.5%
Big Data	28.7%
Machine Learning	26.2%
Apache Hadoop	23.9%
Java	21.2%
Business Intelligence	21.2%
Tableau	18.9%
Data Analysis	18.8%
Data Warehousing	16.2%
Extraction Transformation and Loading	15.4%
SAS	13.8%
Project Management	13.6%
Apache Hive	13.6%
R	11.3%
Scala	11.1%
Software Development	11.1%
Data Mining	10.9%
Data Visualization	10.3%
Data Modeling	10.0%
Economics	9.7%

Table 2. General Business Skills Required for Big Data Titled Jobs (January to December 2018)

Skills	Job Postings
Communication Skills	38.2%
Teamwork/Collaboration	28.6%
Research	28.5%
Problem Solving	23.5%
Microsoft Excel	16.8%

Table 3. Big Data Titled Jobs by Cluster, National and Regional (January to December 2018)

Cluster	U.S. Total Jobs	New Jersey/New York MSA Jobs	Sample New Jersey/New York Employers
Professional, Scientific, Technical Services	39,001	3,458	Deloitte, IBM, Accenture, KPMG, McKinsey, PWC
Finance and Insurance	24,304	3,966	JP Morgan, Citi, Anthem, Goldman Sachs
Information Technology	11,281	1,664	Comcast, Verizon, Bloomberg, ATT
Life Science/Pharma Other Chemicals	3,358	732	J&J, Merck, Pfizer, BMS, Glaxo, Novartis, Regeneron
Advanced Manufacturing*	1,146	87	Merck, J&J, GE
Food Manufacturing	471	45	Mars, Mondelez, IFF, Campbell

* This category is defined by Labor Insight, and includes pharma and other advanced manufacturing

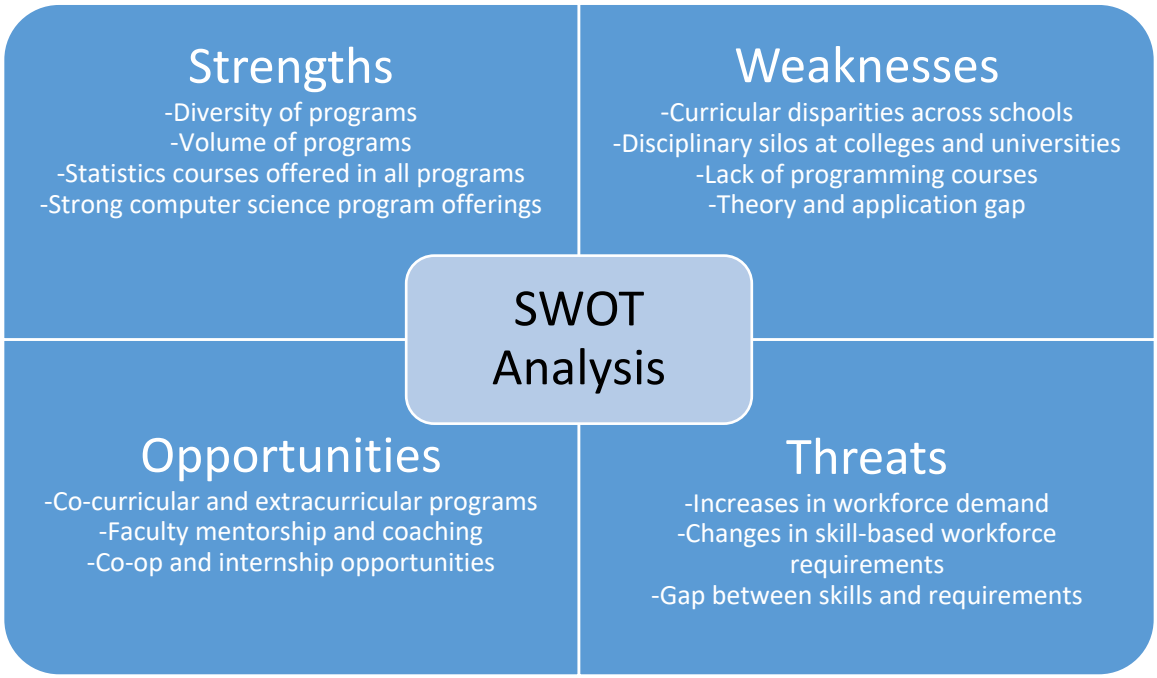
Table 4. Data Science and Business Analytics Program Offerings in Four-Year Colleges and Universities (Undergraduate and Graduate) in New Jersey Region (AY 2019)

	Undergraduate	Graduate
Minor in Data Science or Related Field	85%	65%
Major in Data Science or Related Field	50%	77%
Major in Computer Science with Data Science, Artificial Intelligence, Big Data, or Analytics Focus	100%	65%

Table 5. Big Data Workforce Roadmap for New Jersey

Career Stage	Skill Sets in Demand	Individual (Student and/or Worker) Role	Educator Role	Employer Role
Pre-Career: K-12	• Learning to learn • Growth mindsets and curiosity • Ability to collaborate with others • Ability to communicate	• Coding clubs • Involvement in robotics competitions • Science fair participation • Maker fair participation • Hack-a-thons	• Incorporate data science skills into existing curriculum • Strong math, physics, and science instruction at high school level	• Participate in and/or sponsor fairs, Hack-a-thons • On-site summer events
Pre-Career: Community College Four-Year College County Tech Academies	• Math and statistical foundations • Computational foundations • Data management, curation, data visualization, data modeling, workflow, and reproducibility • Communication/teamwork • Ethical problem solving	• Explore career pathways • Join professional societies and student clubs to learn through engagement • Understand certifications and where to get them • Take courses in data analytics/programming • Participate in Hack-a-thons, internships/ externships • Take a capstone course	• Require at least one data science course for all students • Offer capstone courses to allow for student interactions from different disciplines • Leverage advanced computing technologies to improve education and training • Outcomes-based education using real-world examples demonstrating data usage and case studies	• Internships/externships for students using real-world, structured and unstructured data • Sponsor Hack-a-thons • Provide mentors
Entry-Level Career	Dependent on job specifications	• Be willing to apply skill sets and learn new ones • Develop personal insights for chosen career and "pivot" as needed.	• Educator becomes a coach that students connect with through LinkedIn and University alumni network	• Individual Development Plan for employees • Experiential and digitally based training for new hires
Mid-Career	Dependent on job specifications	• Demonstrate "learnability" • Be aware of tech trends and up-skill as needed	• Commitment to programs that support lifelong learning	• Financially support up-skilling • Encourage employees to participate in student mentoring
Advanced Career	Dependent on job specifications	• Capture and share on-the-job knowledge and experience	• Continued relationship building with alumni	• Encourage employees to volunteer and teach in educational programs
Post-Career	NA	• Mentor students or start-ups • Teach at local college	• Tap knowledge of retirees as start-up mentors; invite them to guest lecture; etc.	NA

Figure 1. SWOT Analysis of Data Science and Business Analytics Programs in New Jersey



References

Aasheim, C. L., Rutner, P., Williams, S., & Gardiner, A. (2014). Big data analytics and data science undergraduate degree programs. *2014 Annual Meeting of the Decision Sciences Institute Proceedings* (2014) pp. 338-359.

Arboleda, P. (2018). *Consumers want connected medical devices, but demand for digital experts could put further strains on the talent pool for medtech*. Retrieved January 16, 2019, from https://blogs.deloitte.com/centerforhealthsolutions/consumers-connected-medical-devices-demand-digital-experts-strains-talent-pool-medtech/

Berinato, S. (2019). Data science and the art of persuasion. *Harvard Business Review*, Jan./Feb., 126-137.

BioNJ. (2018). *The New Jersey biopharma industry: A prescription for growth*. Trenton, NJ: Author. Retrieved from https://bionj.org/wp-content/uploads/2018/02/BioNJ-Full-White-Paper-012918.pdf

Bowley, R. (2018). LinkedIn data reveals the most promising jobs and in-demand skills of 2018. Retrieved January 25, 2019, from https://blog.linkedin.com/2018/january/11/linkedin-data-reveals-the-most-promising-jobs-and-in-demand-skills-2018

Burning Glass Technologies. (2019). *The hybrid job economy*. Boston: Author. Retrieved from https://www.burning-glass.com/wp-content/uploads/hybrid_jobs_2019_final.pdf

Catlin, T., Scanlan, J., & Willmott, P. (2015). *Raising your digital quotient.* New York: McKinsey & Co. Retrieved from https://www.mckinsey.com/business-functions/strategy-and-corporate-finance/our-insights/raising-your-digital-quotient

Cegielski, C. G., & Jones-Farmer, L. A. (2016). Knowledge, skills, and abilities for entry-level business analytics positions: A multi-method study. *Decision Sciences Journal of*

Innovative Education, *14*(1), 91-118.

Cleary, J. (2017). Jobs and skills in demand: Cybersecurity. Retrieved January 25, 2019, from https://mbs.rutgers.edu/articles/job-skills-demand-cybersecurity

Donovan, F. (2017). Big data analytics is expected to have the biggest technology impact on the pharmaceutical industry in 2019, according to a survey of pharm professionals by GlobalData. *HIT Infrastructure*. Danvers, MA: Xtelligent Media, LLC**.** Retrieved from https://hitinfrastructure.com/news/big-data-analytics-to-have-major-impact-on-pharma-next-year

Giles, M. (2018). AI for cybersecurity is a hot new thing — and a dangerous gamble. *MIT Technology Review*. Retrieved from https://www.technologyreview.com/s/611860/ai-for-cybersecurity-is-a-hot-new-thing-and-a-dangerous-gamble/

Google. (2018). Computer learns to detect skin cancer more accurately than doctors. *The Guardian*, *29*.

Irani, Z., Sharif, A. M., Lee, H., Aktas, E., Topaloglu, Z., van't Wout, T., & Huda, S. (2018). Managing food security through food waste and loss: Small data to big data. *Computers & Operations Research*, *98*, 367-383

Kent, J. (2018). *Deep learning, big data fuel medical device for predicting seizures*. Retrieved from https://healthitanalytics.com/news/deep-learning-big-data-fuel-medical-device-for-predicting-seizures

Krause, R. (2018). AI companies race to get upper hand in cybersecurity. *Investor's Business Daily*. Retrieved from https://www.investors.com/news/technology/ai-companies-artificial-intelligence-cybersecurity/

LaborInsight. (2018). *Labor insight: Real time market data*. Boston: Burning Glass

Technologies. Retrieved January 25, 2019, from https://www.burning-glass.com/products/labor-insight/

Lee, J., Davari, H., Singh, J., & Pandhare, V. (2018). Industrial artificial intelligence for industry 4.0-based manufacturing systems. *Manufacturing Letters*, *18*, 20-23.

Loucks, J., Davenport, T., & Schatsky, D. (2018). State of AI in the Enterprise, 2nd ed. *Deloitte Insights*. Retrieved from https://www2.deloitte.com/insights/us/en/focus/cognitive-technologies/state-of-ai-and-intelligent-automation-in-business-survey.html

McKinsey Global Institute. (2018). *Skill shift: Automation and the future of the workforce* (Discussion Paper, May 2018). New York: Author. Retrieved from https://www.mckinsey.com/~/media/McKinsey/Featured Insights/Future of Organizations/Skill shift Automation and the future of the workforce/MGI-Skill-Shift-Automation-and-future-of-the-workforce-May-2018.ashx

MITRE. (2017). *Artificial intelligence for health and health care*. *JASON Report* (Vol. 7508). Retrieved from https://www.healthit.gov/sites/default/files/jsr-17-task-002_aiforhealthandhealthcare12122017.pdf

Monegain, B. (2016). Amazon, Google, IBM, Microsoft join forces with MIT and Harvard on cloud-based genome analysis toolkit. *Healthcare IT News*. Retrieved from https://www.healthcareitnews.com/news/amazon-google-ibm-microsoft-join-forces-mit-and-harvard-cloud-based-genome-analysis-toolkit

National Academies of Sciences, Engineering, and Medicine. (2018). *Data science for undergraduates: opportunities and options*. Washington, D.C.: National Academies Press.

New Jersey Business Magazine. (2017). CyberSeek details supply and demand of cybersecurity

workers in NJ. *New Jersey Business*. Retrieved from https://njbmagazine.com/njb-news-now/cyberseek-details-supply-demand-cybersecurity-workers-nj/

New Jersey Department of Labor and Workforce Development. (2018). *New Jersey's advanced manufacturing industry cluster*. Trenton, NJ: Author.

New Jersey Department of Labor and Workforce Development. (2019). *Industry & occupational employment projections*. Trenton, NJ: Author. Retrieved from https://www.nj.gov/labor/lpa/employ/indoccpj/indoccpj_index.html

New Jersey Economic Development Authority. (2018). *The state of innovation: Building a stronger and fairer New Jersey*. Trenton, NJ: Author. Retrieved from https://www.njeda.com/pdfs/StrongerAndFairerNewJerseyEconomyReport.aspx.

Oracle. (2015). *Improving pharmaceutical & life sciences performance with big data*. Oracle Enterprise Architecture White Papers. Redwood City, CA: Author. Retrieved from http://www.oracle.com/us/technologies/big-data/big-data-pharmaceutical-2419817.pdf

Perry, T. (2018). *Desperate for data scientists*. New York: IEEE Spectrum. Retrieved from https://spectrum.ieee.org/view-from-the-valley/at-work/tech-careers/desperate-for-data-scientists

Perry, T. (2019). *Demand and salaries for data Scientists continue to climb*. New York: IEEE Spectrum. Retrieved from https://spectrum.ieee.org/view-from-the-valley/at-work/tech-careers/demand-and-salaries-for-data-scientists-continue-to-climb

Rao, A. S., Verweij, G., & Cameron, E. (2017). Sizing the prize: What's the real value of AI for your business and how can you capitalise? *PwC*, 32. Retrieved from https://www.pwc.com/gx/en/issues/analytics/assets/pwc-ai-analysis-sizing-the-prize-report.pdf

Sachs, M. (2018). *Energy storage and big data* [PowerPoint slides].

Select Committee on Artificial Intelligence. (2018). *AI in the UK: Ready, willing and able?* (Vol. 60). London: House of Lords.

Shoham, Y., Perrault, R., Brynjolfsson, E., Clark, J., Manyika, J., Niebles, J. C., … Bauer, Z. (2018). The AI index 2018 annual report. *AI index steering committee, human-centered AI initiative*. Stanford, CA: Stanford University. Retrieved from http://cdn.aiindex.org/2018/AI Index 2018 Annual Report.pdf

Sigelman, M. (2018). *Skills, not jobs: How higher ed can capitalize on big data's disruptive moment* [PowerPoint slides].

Sikos, L. F. (2018). *AI in cybersecurity* (Vol. 151). New York: Springer.

Silver, D., Bernis, K., & Weston, S. (2018). *Is analytics always the same? An industry-based exploration of skills needed for analytics jobs*. Retrieved January 25, 2019, from https://mbs.rutgers.edu/articles/analytics-always-same-industry-based-exploration-skills-needed-analytics-jobs

Stone, P., Brooks, R., Brynjolfsson, E., Calo, R., Etzioni, O., Hager, G., … Kraus, S. (2016). Artificial intelligence and life in 2030. *One Hundred Year Study on Artificial Intelligence: Report of the 2015-2016 Study Panel*. Stanford, CA: Stanford University.

Tang, R., & Sae-Lim, W. (2016). Data science programs in U.S. higher education: An exploratory content analysis of program description, curriculum structure, and course focus. *Education for Information*, *32*(3), 269-290.

23andMe. (2018). *GSK and 23andMe sign agreement to leverage genetic insights for the development of novel medicines*. Retrieved January 25, 2019, from https://mediacenter.23andme.com/press-releases/gsk-and-23andme-sign-agreement-to-

leverage-genetic-insights-for-the-development-of-novel-medicines/

The White House Office of Science and Technology Policy. (2018). *2018 White House summit on artificial intelligence*. Washington, D.C.: Author.

Weber, L. (2019, January 22). The hybrid skills that tomorrow's jobs will require. *Wall Street Journal*. Retrieved from https://www.wsj.com/articles/the-hybrid-skills-that-tomorrows-jobs-will-require-11547994266

Wixom, B., Ariyachandra, T., Douglas, D. E., Goul, M., Gupta, B., Iyer, L. S., … Turetken, O. (2014). The current state of business intelligence in academia: The arrival of big data. *CAIS*, *34*, 1.

World Economic Forum. (2018). *The future of jobs report 2018.* Cologny-Geneva, Switzerland: Author. Retrieved from http://www3.weforum.org/docs/WEF_Future_of_Jobs_2018.pdf

Methodology for Sourcing Burning Glass Data

The data about skills came from Labor Insight from Burning Glass (LaborInsight, 2018). The study team created a query for jobs posted in analytics between January and December 2018. The positions were selected based on having a job title containing the words: data scientist, big data, machine learning, data engineer, business intelligence or analytics, or any combination. Study team members were specifically looking for postings that requested a Bachelor's degree or higher. In this report, at times state data (New Jersey-specific jobs) are reported and at other times the MSA was used to define the New Jersey/New York region. MSA is a geographical region with a high population density at its core and economic ties throughout the area. Three MSAs compose the study's region: Philadelphia-Camden-Wilmington, New York-Newark-Jersey City, and Trenton.

While this analysis can show trends in the job market, there are limitations. Only jobs advertised online are included. The unstructured nature of job ads can make it difficult for the system to identify individual pieces of information effectively in some cases. While Labor Insight breaks up the job description into fields for analysis, inconsistency in the formatting of job descriptions and industry-specific terminology or titles may result in the inclusion of some irrelevant jobs. Furthermore, while Labor Insight attempts to remove duplications from recruiters, duplications still exist.

End Notes

[1] The New Jersey Economic Plan identifies the following as high-wage, high-growth industries that will be targeted for support: life sciences, information technology and high technology, clean energy, advanced manufacturing, advanced transportation and logistics (including aviation), finance and insurance, food and beverage, and film and digital media (New Jersey Economic Development Authority, 2018, p. 42).

[2] Please see explanation of methodology used to obtain these data.

[3] Ibid.

[4] This does not include people who may have big data jobs in New York City, but live in New Jersey.

[5] These clusters are based on Industry sector North American Industry Classification System. Thirty-three percent of the jobs are unclassified and are not included in this table.

[6] Certified information systems auditing certification is earned through a qualification with ISACA.org that allows IT professionals to maintain updated professional skills in compliance, data security standards and audit controls for business and information systems.

http://www.isaca.org/Certification/CISA-Certified-Information-Systems-Auditor/Pages/default.aspx.