

Delphi study sets out to identify essential cognitive, interpersonal, and intrapersonal competencies needed for success in data analytics.

BY ORLI WEISER, YORAM M. KALMAN, CARMEL KENT, AND GILAD RAVID

65 Competencies: Which Ones Should Your Data Analytics Experts Have?

ANALYZING DATA IS NOW essential to success in education, employment, and other areas of activity in the knowledge society. Even though several frameworks describe the competencies and skills needed to meet current and future challenges, no data analytics competency framework exists to describe the importance of specific skills to succeed in data analytics assignments. In this article, we explore which competencies are required for effective data analytics by applying the Delphi technique and exploring the opinions of data analytics experts.

Our results present a list of cognitive, intrapersonal, and interpersonal competencies, voted up by a consensus panel of experts in the field. Focusing on the three categories of essential competencies will

help to better prepare students and employees for existing and future roles as citizens, employees, managers, parents, and volunteers. We urge policy-makers, academic institutions, and educators to establish programs and reassess existing curricula and materials to support the development of these essential competencies.

The Evolving Field of Data Analytics

Analyzing data has become mandatory for making successful decisions and accomplishing tasks, not only in the workplace but also across many contexts related to health, education, leisure, citizenship, and more.²⁸ The demand for professionals in the rapidly growing field of data analytics keeps rising.^{2,22} Hence, identifying, evaluating, and teaching data analytics competencies is an important goal for academic institutions and organizations in every sector. However, we must understand what those competencies are, and how organizations can identify and measure the competencies of candidates as they pertain to data analytics. A review of the literature reveals discrepancies between contemporary conceptualizations of the skills and capabilities required for data analytics.²⁹

The importance of data. Technological developments, as well as the diffusion of smart devices, have resulted

>> key insights

- **Data analytics is the process of examining data to draw conclusions.**
- **Data analytics is an important tool for making everyday organizational decisions as well as for policy-based decisions, such as those used to manage the COVID-19 pandemic.**
- **In this Delphi study, we set out to identify the essential competencies required to effectively use data analytics to prepare for current and future roles, challenges, and responsibilities as educators, managers, policymakers, citizens, and parents.**
- **We identified cognitive, interpersonal, and intrapersonal competencies needed for successful data analytics.**

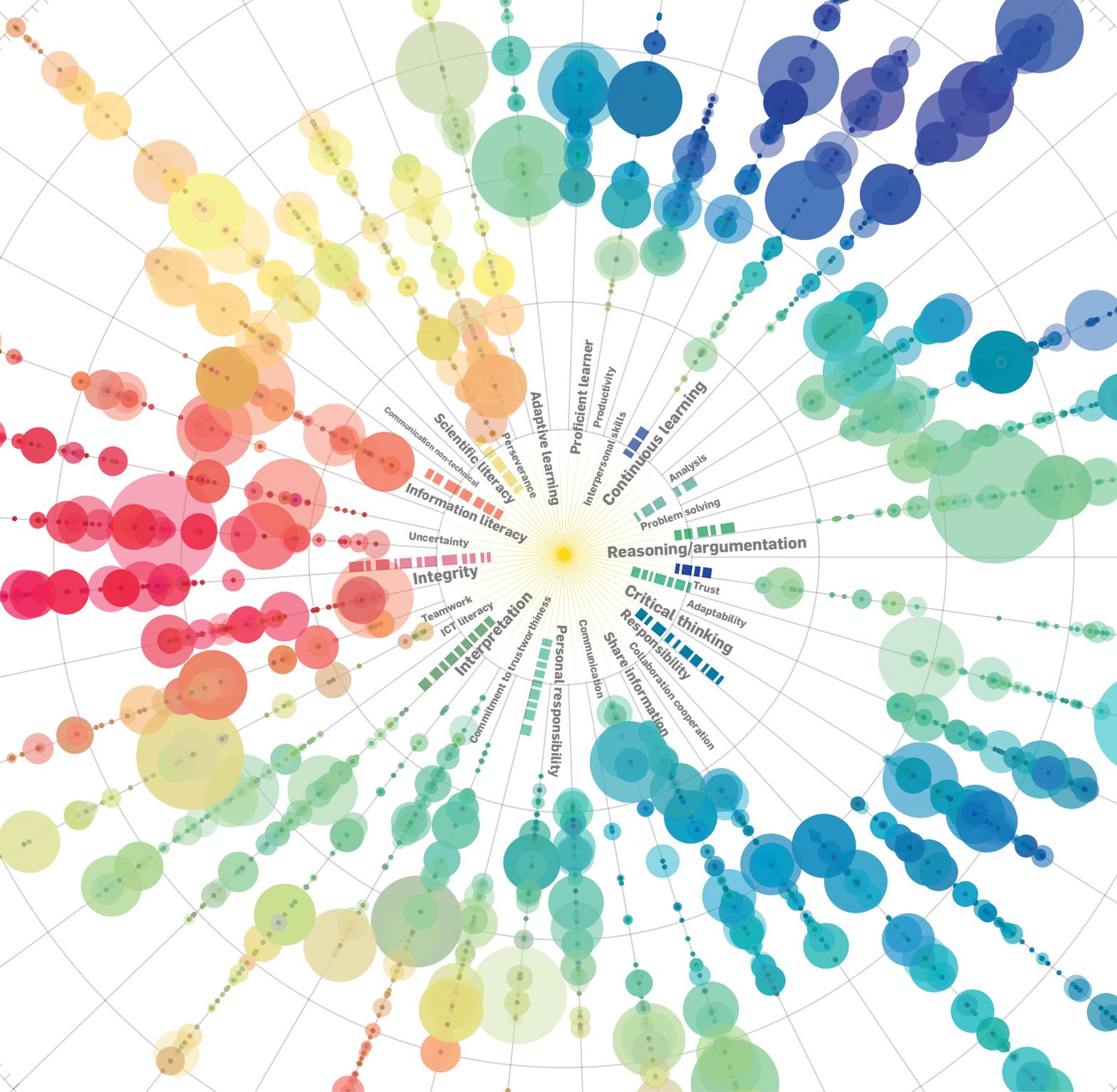


IMAGE BY ANDRĀJ BORYS ASSOCIATES, USING SHUTTERSTOCK

in large volumes of data that are generated at an unprecedented rate. The formats of data and the way humans interact with it are rapidly changing, requiring people to become “data literate.” They need to become active data explorers who can plan for, acquire, manage, analyze, and infer insights from data.¹⁰ At the same time, technological developments also enable organizations to efficiently manage huge data sets and to store them for long periods. Hence, data becomes a vital factor for both individual and organi-

zational success. The term “big data” was coined by Roger Magoulas from O’Reilly Media in 2005. It refers to a wide range of large datasets almost impossible to manage and process using traditional data management tools—due to their complexity as well as their size.¹³ Data can exist in the form of documents, email, text messages, audio, images, video, graphics data, and more.^{12,30}

Data analytics is the process of examining data to draw conclusions and make better decisions.^{26,29} Orga-

nizations, therefore, use data analytics techniques to make better organizational decisions. This process might, for example, boost business performance by increasing revenue, improving operational efficiency, or optimizing marketing campaigns. Integrated data can support not only academic research and organizational decisions, but also the analytics requirements of government itself. Data that government agencies collect and use to run their programs provides policymakers with new sources of facts for bench-

marking goals and measuring the successes and shortcomings of existing and future programs.¹⁵

The profession of data analytics. In the past, data analytics tasks were performed by a limited number of professionals: scientists, researchers, economists, etc. During the last decade, as developments in science and technology have become widely applied in various sectors of the economy, employer demand for data analytics expertise rose and often resulted in new occupations being defined.²³ Analysts in the organization who work with the data are key to success in decision-support data-management tasks.³⁶ Data scientists and other professionals skilled in working with large quantities of data are critical to successfully using big data.²⁴ Thus, the demand for professionals who can support the work keeps rising.^{1,11,22} The Occupational Information Network (O*NET) database of occupation descriptions is a program sponsored by the U.S. Department of Labor.²³ The O*NET-SOC 2019 taxonomy, which includes the most up-to-date O*NET list, contains descriptions of 1,016 occupations, covering the entire U.S. economy. Many of the new data occupations—for example, data scientists and digital forensics analysts—were added to the taxonomy only recently, in 2018 and 2019,³² further evidence of the field’s dynamic nature and uncertainty.

The data analytics process comprises many tasks,²⁰ which are usually performed by an expert. The extract, transform, and load (ETL) process starts with collecting data from many sources and combining them to transform the different types of data into a common format. The next stage is cleansing the data. In this process, duplicate and irrelevant records are deleted, missing values are filled, and the data is checked for errors. When the data is ready for analysis, it is uploaded into the analytics system, where it can serve as the basis for a model that can be trained to forecast future information. The data can also be presented in data visualization systems such as dashboards, graphs, and charts.

Professional competencies. What is the difference between a skill and a



Identifying, evaluating, and teaching data analytics competencies is an important goal for academic institutions and organizations in every sector.



competency? A skill is a specific, learnable ability that one needs to perform a given job well.²⁷ A competency is more than just knowledge and skills. It involves the ability to meet complex demands by drawing on and mobilizing psychosocial resources, including skills and attitudes, in a particular context.²⁷ In this research, we used the term “competencies” rather than “skills” to reflect our view that skills and knowledge are intertwined, as suggested by Pellegrino and Hilton,²⁸ based on OECD terminology.²⁷

Professional competencies are skills, knowledge, and attributes that refer to the ability to perform effectively within a professional role. They are usually the competencies one must present during a job interview. In today’s economy, however, workers must be more collaborative and creative in their problem-solving techniques. While “hard” technical skills associated with programming—for instance, programming languages and databases—remain a prerequisite for new hires, the industry also wants workers who can demonstrate a range of so-called “soft” skills, as well as the resiliency and flexibility to work on a range of tasks. Thus, the existing STEM education gap, which is relevant to the data analytics profession, represents a deficit not only in the hard cognitive skills associated with science and engineering, but also in the soft interpersonal and intrapersonal skills linked with effective communication, collaboration, and adaptability.⁴

Existing competency frameworks.

A competency framework consists of a set of specific competencies, bound together in an integrated approach.²⁷ A competency model is a descriptive tool that identifies the skills, knowledge, personal characteristics, and behaviors needed to effectively perform a role in the organization.²¹ Since there is no specific data analytics competency framework, policymakers and educators often refer to other sets of skills such as “21st-century skills,” “deeper-learning skills,” or “higher-order-thinking skills.” Competency frameworks are used for hiring new employees, training them, evaluating their performance, promoting them, and developing their careers, as well

as to support organizational change.⁵

Data analytics competencies are an individual’s personal characteristics that may influence how that person approaches data analytics tasks and acquires data-relevant knowledge and skills. Though several frameworks are called on to describe the competencies and skills that children and adults require to meet the world’s current and future complex challenges, to the best of our knowledge, no specific data analytics competency framework exists. In this study, we take one more step toward closing this gap by compiling an expert consensus around the required competencies for data analytics practitioners. The study is based on a profession-related framework,³⁴ education, and other life context-related frameworks.^{10,28} By understanding the ways that organizations use competency frameworks, higher education institutions can develop curriculums to enhance student development.²¹

*AIS IS 2010 Curriculum Guidelines.*³⁴ Information systems (IS) are complex systems requiring both technical and organizational expertise for design, development, and management. The availability of curriculum models enables local academic units to maintain academic programs that are consistent with regional, national, or global employment needs and with the common body of knowledge of the IS field. The curriculum is designed to educate and prepare graduates to enter the workforce by equipping them with the knowledge and skills specified in three categories: IS-specific knowledge and skills, foundational knowledge and skills, and domain fundamentals.

The “IS 2010” report, a collaborative effort between the Association for Computing Machinery (ACM) and the Association for Information Systems (AIS), is the latest output from model curriculum work for information systems that began in the early 1970s. It is grounded in the expected requirements of the industry, represents the views of organizations employing graduates, and is supported by other IS-related organizations. The report identifies prerequisite skills needed by all students in basic personal-productivity software (word processing,

email, Web browsing, spreadsheet modeling, etc.) and lists skills that IS professionals must possess (critical thinking, collaboration, effective communication, persistence, flexibility, curiosity, creativity, etc.).

*National Research Council (NRC) report—Education for life and work: Developing transferable knowledge and skills in the 21st century.*²⁸ The National Research Council (NRC) is the principal operating agency for both the National Academy of Sciences and the National Academy of Engineering, providing services to the government, the public, and the scientific and engineering communities. The council connects the broad science and technology community with the academies’ purposes of furthering knowledge.

In its report, the committee addressed questions about the terms “21st-century skills” and “deeper learning,” as well as about their educational and social implications. The committee examined the evidence for the importance of various types of competencies for success in education, work, health, and other life contexts, and drew on a large research base in cognitive, developmental, educational, orga-

nizational, and social psychology and economics.

To organize the various terms for 21st-century skills and to provide a starting point for further research as to their meaning and value, the committee identified three broad domains of competence—cognitive, interpersonal, and intrapersonal—which represent distinct facets of human thinking. The *cognitive domain* involves reasoning and memory, and it includes competencies such as analysis, problem solving, scientific literacy, and creativity. The *intrapersonal domain* deals with the capacity to manage one’s behavior and emotions to achieve one’s goals. It includes competencies such as perseverance, adaptability, flexibility, self-direction, and the ability to cope with uncertainty. The *interpersonal domain* includes competencies, such as communication and collaboration,²⁸ which are used to express information to others, to interpret others’ messages (both verbal and nonverbal), and to respond appropriately. It involves expressing ideas and interpreting and responding to messages from others.

*Innovating Pedagogy report.*¹⁰ This report is the sixth in a series of annual

Figure 1. The Delphi Study.

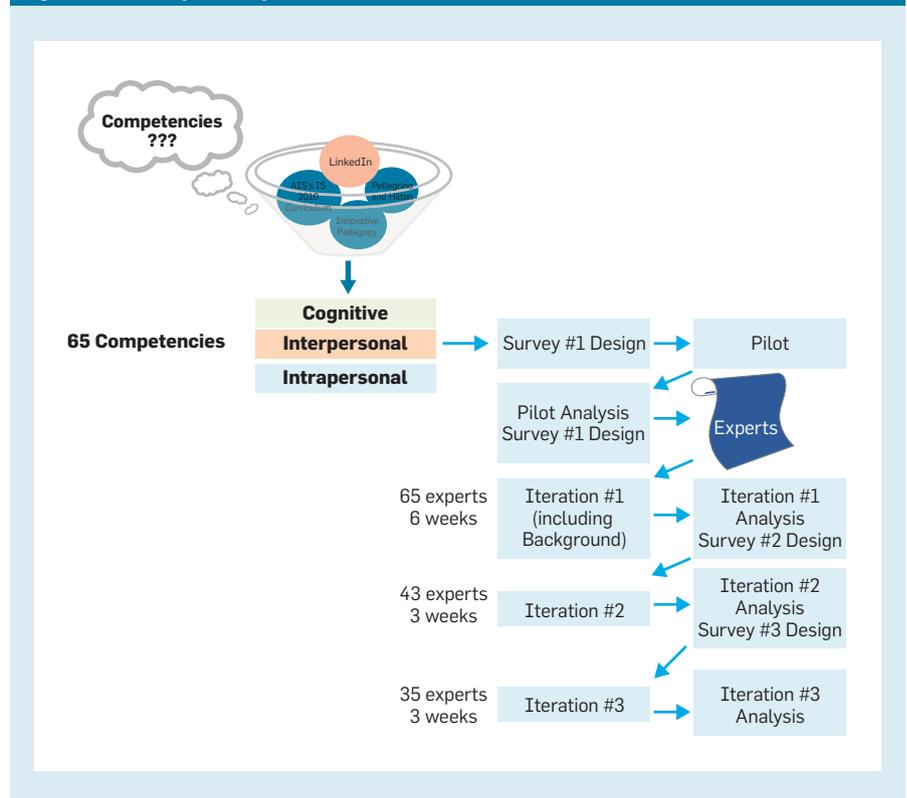


Table 1. Expert criteria and characteristics for the Delphi study.

Criteria	Requirements	Participated (Iteration #1)
Number of experts	Minimum of 50 experts	65 participants
Place of work and research	Minimum of 25 experts (worked/researched in more than one country)	Worked/researched in more than one country : 33 participants Worked/researched in only one country : 32 participants (of which 26 are in Israel)
Field of work	A diverse selection of industry, governmental and non-governmental organizational, and academia. Minimum of eight experts for each sector	Industry : 31 participants Governmental/non-governmental non-profit organizations : 14 participants Academia : 28 participants
Gender	At least 15 experts for each gender	Male : 44 participants Female : 21 participants
Education	Relevant for the domain of expertise. No formal degree required.	Relevant for the domain of expertise: 65 participants
Professional experience	Minimum three years of experience in data analytics. The average experience in data analytics of all participants will be at least five years. At least 10 experts will have some managerial/decision-level experience in the field.	Expertise Level Expert : 31 participants Knowledgeable : 33 participants Novice : 1 participant Participant Years of Expertise Less than 3 years : 4 participants 3 years : 8 participants 4–7 years : 20 participants 8–10 years : 8 participants 11–15 years : 6 participants Over 15 years : 18 participants
The domain of expertise in data analytics/data science	Relevant to the study. At least one expert from each phase of the CRISP-DM model	Domains of Expertise Data science : 26 participants Data engineering : 16 participants Data analytics : 34 participants Statistics : 11 participants Other : 27 participants CRISP-DM Model Phases Business understanding : 45 participants Data understanding : 64 participants Data preparation : 55 participants Modeling : 47 participants Evaluation : 47 participants Deployment : 37 participants Other : 4 participants

reports on innovations in teaching, learning, and assessment. While it is not the most up-to-date report, it is the most relevant for the current research, since it introduces 10 pedagogies that either already influence educational practice or offer opportunities for the future. The report is the result of a collaboration between researchers at the Institute of Educational Technology in The Open University, U.K., and the Learning In a NetworKed Society (LINKS) Israeli Center of Research Excellence (I-CORE).

The 2017 report lists crucial skills that are required to become proficient learners and citizens: problem solving, evidence evaluation, and making sense of complex information from various

sources. The report also highlights STEM topics that can develop these skills and addresses current demands for STEM-skilled employees across job sectors. Therefore, one of the recommendations for today’s students, who need to be prepared for a data-driven society, is to learn to work and think with data.

The Delphi Technique

One of the ways to close knowledge gaps is to ask experts for their opinions, predictions, and estimations about the future impact of current technological changes. The Delphi technique is a known structured process that uses a series of questionnaire-based iterations to gather

information from a chosen panel of respondents within their domain of expertise until group consensus is achieved.⁸ This technique applies to goal setting, policy investigation, and predicting future events.¹⁶ Thus, the Delphi technique is used for acquiring knowledge about new and emerging fields in education,⁶ work,¹⁸ and health¹⁴ information systems,³⁵ as well as other life contexts—for example, Baumer et al.³ In the current research, this technique is used to identify which competencies are essential to meeting future challenges of the big data era—that is, data analytics competencies.

Mapping the competencies. The first stage of the Delphi study was a literature review of contemporary skillsets. We mapped the competencies found in the described literature and combined the lists (see a schematic, high-level view of the Delphi study in Figure 1). Next, we integrated this list with a list of competencies appearing in recruitment messages for data analytics-related jobs posted on the career-oriented LinkedIn website from January 2018 through June 2018. The recruitment messages we analyzed contained the phrases such as “data analytics,” “data analyst,” “data scientist,” “data engineer,” and “data mining.”

The analysis was performed iteratively by date until no new competencies were found. We then manually classified the resulting list of competencies into five categories: cognitive competencies, intrapersonal competencies, interpersonal competencies, education, and technologies. We compared this new list to the original competencies list that was prepared based on the described literature, and added any new competencies found in the LinkedIn recruitment messages to the original list. The consolidated list includes 65 competencies (See Appendix at <https://dl.acm.org/doi/10.1145/3467018>).

Developing a survey of competencies. The second stage of the study was the development of a survey, which included the consolidated list and our three competency categories: cognitive, intrapersonal, and interpersonal. Each competency in the survey was accompanied by a definition based on several resources: The *Oxford English Dictionary*, the *Merriam-Webster dic-*

tionary, and Wikipedia. These definitions are also presented in the online Appendix. The survey asked participants to classify the importance of each competency for working in data analytics, with the following possible responses: “Essential competency,” “Important competency,” “Less important competency,” “Not an important competency,” “I do not know the importance,” and “I do not understand this item.” The items were followed by open text questions, in which the experts were able to elaborate on their responses, as well as add their opinions and suggestions.

Looking for consensus. Following the survey development, 200 experts from industry, government and non-government organizations, and academia from around the world were invited to participate in the study. They were chosen according to predetermined criteria (see Table 1) to ensure their expertise in data analytics. The criteria included the type of employer they work for, their field of work, the international nature of their work, gender, education level, professional experience, domain expertise, and their areas of data analytics expertise, according to the CRISP-DM model phases.³¹ The experts were given one six-week interval and then two three-week intervals to complete each survey iteration. Three weeks separated each of the survey time frames, allowing the research team to analyze results and create the next survey. Typical in this type of study, as time progressed, participation requirements increased and so did attrition rates. Sixty-five experts accepted the invitation and responded to the first survey, and 35 ultimately completed the third survey.

After analyzing replies to the first survey, we used the interquartile range (IQR) method to set the consensus level according to the relatively strict level of 75% of replies.^{16,19} Since participant feedback on interim findings is one of the most important characteristics of the Delphi technique,³⁵ the second survey’s main goal was to reflect the opinions identified in the first survey and allow the experts to respond to these initial findings. Accordingly, the second survey was developed based on the consensus found in the first survey. It included

two parts. In the first, we presented two updated definitions of competencies that were not clear in the first survey, as well as a list of 32 competencies that didn’t reach the consensus level. At least 15% of the participants in the first survey were unclear about the definitions of “Executive function,” “Self-regulation type 1,” and “Self-regulation type 2.” After consideration, we replaced these three competencies with one, “Self-regulation,” and added a new competency: “Perceived self-efficacy.” We asked the experts to reconsider the importance of these two competencies in addition to the 32 competencies that did not achieve consensus. Participants were asked to classify the importance of the competencies for working in data analytics, with the same possible responses from the first survey.

In the second part of this survey, we presented the experts with the 30 competencies that achieved consensus level, defined as more than 75% of the experts agreeing that a particular competency is important for data analytics,

or more than 75% agreed that it is not. We then offered the experts a chance to share their thoughts and suggestions on this list. Forty-three experts responded to the second survey, with many of them expressing their satisfaction with the results of the first iteration and the relevance of the findings (see Table 2). With the help of the experts’ comments, we were able to ensure they followed instructions and responded to the questionnaire in the context of data analytics. Following the second survey, the consensus calculation was repeated. Two additional competencies reached a consensus, leading to a minor adjustment of the findings and the creation of a third survey, which described the final list of competencies that reached consensus and again asked the experts for their opinion about the list.

Results—The Most Important Competencies Your Data Analytics Experts Should Have

The results of the first and the second surveys revealed 34 competencies that

Table 2. Iteration #2—Experts’ opinions (selected).

Q1: What do you think about this list of important competencies?

Expert ID	Answer
14	“These are all important competencies, and some of them are essential. However, nobody is perfect, so it may be hard to find people who excel in all of them. Hence, beyond the necessary cognitive skills, teamwork, collaboration, and continuous learning, are essential.”
33	“This is a very good summary of the skills needed for a data analyst to do his or her jobs well.”
8	“They are important not only to the data analyst . In general I think that most of the competencies you listed are important for any high-functioning person in many fields.”
39	“Depending on the context, most of these themes, if not all, are very important. To add, in my opinion, the need for inductive reasoning in data analytics .”
12	“It’s a good wish list.”
7	“Fairly wide and comprehensive, just like the job title ‘ data analysis .’ It might be that forcing choice between qualities in more of a zero-sum setting could yield more insights.”
65	“I am surprised to see that teamwork and communication are so low down on the list as the best data analytics work is actually done through collaboration in my opinion.”

Q2: Do you have anything else to add about the questionnaire or about the topic of data analytics?

Expert ID	Answer
42	“Nothing to add. However, the problem with this list is that many of the competencies are not very trainable. Many are shaped by genetics and early life experience. By relying on these competencies you are selecting for upper middle class, well-socialized people who, by definition, are part of dominant culture that has defined the competencies. I may be very wrong in my assumptions and would be thrilled to be disproven.”
23	“I will be interesting to see the results of this research.”

Table 3. Important competencies for data analytics.

Consensus	Cognitive	Intrapersonal	Interpersonal	Responses (N)	Important/ Essential (N)	Consensus
Iteration #1						
Analysis	•			64	62	97%
Continuous learning		•		65	61	94%
Problem solving	•			65	60	92%
Reasoning/ argumentation	•			64	58	91%
Critical thinking	•			64	57	89%
Trust			•	64	57	89%
Intellectual interest		•		65	56	86%
Responsibility		•		65	56	86%
Collaboration cooperation			•	65	56	86%
Personal responsibility		•		62	53	85%
Professionalism		•		62	53	85%
Share information			•	60	51	85%
Communication			•	64	54	84%
Commitment to trustworthiness			•	64	54	84%
Interpretation	•			62	52	84%
Teamwork			•	65	54	83%
Communication non- technical			•	65	54	83%
Integrity		•		64	53	83%
Uncertainty		•		64	53	83%
Scientific literacy	•			65	53	82%
Information literacy	•			62	50	81%
Proficient learner		•		62	50	81%
ICT literacy	•			62	49	79%
Productivity		•		63	49	78%
Perseverance		•		62	48	77%
Adaptive learning	•			60	46	77%
Adaptability		•		61	46	75%
Interpersonal skills			•	65	49	75%
Iteration #2						
Oral communication	•			43	33	77%
Proactive		•		43	35	81%

For a detailed explanation of each competency see the online Appendix. The consensus was calculated by dividing the number of Important/Essential responses by the total number of responses.

achieved a consensus of at least 75%. In the first iteration, experts agreed that 28 competencies are important (see Table 3) and four are not (see Table 4). The second survey reflected the opinions revealed in the first survey. It described the competencies that reached a consensus and asked the experts to reconsider the list of competencies that did not reach a consensus. Two additional competencies reached a consensus in the second survey on

top of the first one (see Table 3: Oral communication, 77%; Proactive, 81%).

While the list of non-important competencies for data analytics (see Table 4) only contains four out of 18 (22%) intrapersonal competencies, the list of important competencies for data analytics (see Table 3) includes 10 out of 19 (53%) cognitive competencies, 12 out of 18 (67%) intrapersonal competencies, and eight out of 18 (44%) interpersonal competencies. The experts

chose mostly cognitive competencies as important for data analytics.

The research team expected the list of competencies that did reach consensus to include all competencies that are part of more than one framework on which the study was based. Surprisingly, the list did not include the “empathy” or “negotiation” competencies, which are part of all three frameworks. Neither were mentioned on LinkedIn, which might explain their absence from the most important competencies list. Nor does the list include many competencies that exist in two frameworks, such as “creativity” and “flexibility.” Not only do these competencies exist in two frameworks, but they were also listed on LinkedIn. This gap between experts’ expectations and actual survey results emphasizes the diversity of opinions on this topic and the need for a specific data-analytics competency framework to ensure the focus is right.

One of the most surprising findings of this study is that data analytics experts did not choose “ethics” as one of the most important competencies. Nevertheless, the results did include the “integrity” competency. Ethical issues are constantly raised in new industrial, educational, and governmental data science and artificial intelligence (AI) projects,²⁵ and though only integrity was included in the final list, it makes sense to emphasize the importance of the possibly more comprehensive competency of ethics. Researchers and organizations are facing new challenges that often lie outside their training and comfort zones.³⁷ In addition, our results reflect the consequences of existing and past academic curricula, where the ethics course is not as important as, for instance, programming.⁷ Finally, they also reveal the industry’s attitude toward ethics questions, as reflected in events such as Google’s firing of Timnit Gebru, a former Google ethics engineer who co-authored a paper raising ethical and privacy concerns about the company.¹⁷

To explain these findings and to explore the reason why some competencies were absent from the final list, the research team analyzed the results of the three surveys according to the participants’ affiliation. Affilia-

tion was considered based on a variety of characteristics: national affiliation (place of work and research), field of work, gender, expertise level, years of expertise, managerial experience, domain of expertise, and characterization of data analytics work (CRISP-DM model phase) (Table 1). Differences between the groups were analyzed using cluster analysis and factor analysis methods. A Mann-Whitney U test was run—without correction for familywise error rate due to multiple comparisons—to determine if there were differences in the competency importance scores between the groups.

After correction for familywise error rate due to multiple comparisons, the analysis did not reveal any statistically significant differences, probably due to the large number of comparisons and the relatively small number of participants. Nevertheless, the differences that were detected (presented in Table 5) are interesting and suggest directions for future quantitative research. These results indicated that experts working in academia possibly find competencies such as “critical thinking,” “interpretation,” “social responsibility,” and “ethics” more important than experts working in the industry. Respectively, experts working in the industry possibly find competencies such as “career orientation” and “handling multiple tasks” more important than experts working in academia. When compared to male experts, female experts possibly find the competencies “information literacy” and “being a proficient learner” to be more important. Experts with international experience possibly find that “scientific literacy” is more important compared to experts without international experience, while experts without managerial experience possibly find that “grit” is more important compared to experts with managerial experience.

To summarize, for most of the competencies, distributions of the competency importance scores were similar. Hence, for most of the competencies, there were no significant differences between the different groups.

Conclusions and Future Work

The starting point of this research was several frameworks which list impor-

tant general “21st-century competencies.” The output of this research is a list of the most important competencies to succeed in data analytics assignments, based on a carefully confirmed consensus between data analytics experts. This consensus was strong, and no statistically significant differences were identified between the classifications of different groups of experts.

Interestingly, there was no consensus among the experts that “ethics” is one of the most important competencies for success in data analytics assignments. There was a preliminary indication that opinions on this matter differ in academia versus industry, suggesting further exploration of this issue. Scholars, funders, and regulators should seek opportunities for collaborative, integrative, and innovative

Table 4. Non-important competencies for data analytics.

Consensus	Cognitive	Intrapersonal	Interpersonal	Responses (N)	Important/Essential (N)	Consensus
Iteration #1						
Risk-taking		•		61	14	23%
Physical health		•		55	11	20%
Career orientation		•		63	12	19%
Artistic		•		58	7	12%

Notes: For a detailed explanation of each competency see the online Appendix. The consensus was calculated by dividing the number of Important/Essential responses by the total number of responses.

Table 5. Iteration #1—Differences between groups.

	Field of Work						Sig
	Academia			Industry			
	(N)	Median	Mean Rank	(N)	Median	Mean Rank	
Critical thinking	21	3.00	27.19	24	2.50	19.33	.018
Interpretation	22	3.00	27.20	21	2.00	16.55	.002
Social responsibility	18	1.50	24.06	21	1.00	16.52	.040
Ethics	22	2.00	27.89	23	1.00	18.33	.011
Career orientation	21	1.00	18.50	23	1.00	26.15	.027
Handle multiple tasks	21	1.00	18.90	24	1.00	26.58	.033
	Gender						Sig
	Male			Female			
	(N)	Median	Mean Rank	(N)	Median	Mean Rank	
Information literacy	42	2.00	28.56	20	3.00	37.67	.045
Proficient learner	41	2.00	28.05	21	3.00	38.24	.024
	International Experience						Sig
	Yes			No			
	(N)	Median	Mean Rank	(N)	Median	Mean Rank	
Adaptive Learning	30	2.00	26.00	29	2.00	34.14	.048
Scientific literacy	33	2.00	37.00	31	2.00	27.71	.031
	Managerial Experience						Sig
	Yes			No			
	(N)	Median	Mean Rank	(N)	Median	Mean Rank	
Grit	42	2.00	27.79	18	2.00	36.83	.047

Notes. Median: 0 = Not an important competency; 1 = Less important competency; 2 = Important competency; 3 = Essential competency. Sig: p < .05. After correction for familywise error rate due to multiple comparisons none of these differences were statistically significant

more online

An online-only appendix for this article can be found at <https://dl.acm.org/doi/10.1145/3467018>.

approaches to ensure that big data analytics in science and industry are responsive to foundational ethical concerns.²⁵ We also propose taking a closer look at competencies that were represented in more than one framework but still did not reach a consensus between the experts.

As concepts such as data science, AI, machine learning, deep learning, and distributed and quantum computing gradually enter the mainstream, knowledge workers who originally focused mainly on technology and programing will require a deeper set of competencies and a deeper understanding of the challenges associated with these concepts.³³ Identifying the essential data analytics competencies is an important step towards overcoming this challenge. Yet, as expressed by expert #12 (Table 2), “...this is a good wish list,” and there is still extensive work to be done to develop a workforce that possesses these competencies since, as suggested by expert #42 (Table 2): “...the problem with this list is that many of the competencies are not very trainable. Many are shaped by genetics and early life experience. By relying on these competencies, you are selecting for upper middle-class, well-socialized people who, by definition, are part of dominant culture that has defined the competencies,” and expanding the workforce beyond this group is a key challenge for the educational system, academia, and industry.

This question of trainability should be the focus of future work, but assuming that many of these skills are trainable, we urge policymakers, academic institutions, and educators to establish programs and reassess existing curricula and materials to support the acquisition of these essential competencies by students and employees. Training programs should facilitate the systemic development, implementation, and evaluation of cognitive, intrapersonal, and interpersonal competencies. Focusing on all three categories of essential competencies will help institutions better prepare students and employees for existing and future roles as citizens, employees, managers, and educators.

Acknowledgments

This work was supported by the Open Media and Information Lab at The Open University of Israel (Grant Number 20184), and by two scholarships: one by the Ministry of Science & Technology of Israel (collaboration between academy and industry scholarship, 2020), and one by the office of The President of Israel (Scientific excellence and innovation scholarship, 2020). **C**

References

1. August LinkedIn Workforce Report: Data science skills are in high demand across industries. LinkedIn (2018). <https://news.linkedin.com/2018/8/linkedin-workforce-report-august-2018>.
2. Basu, N. Analytics accelerates into the mainstream. *Forbes Insights* (2017). https://i.forbesimg.com/forbesinsights/d&b_enterprise_analytics/Analytics_Accelerates_Into_Mainstream.pdf.
3. Baumer, E.P., Xu, X., Chu, C., Guha, S., and Gay, G.K. When subjects interpret the data: Social media non-use as a case for adapting the Delphi Method to CSCW. In *Proceedings of the 2017 ACM Conf. on Computer Supported Cooperative Work and Social Computing*, 1527-1543.
4. Burke, Q. and Bailey, C.S. Becoming an 'adaptive' expert. *Communications of the ACM* 63, 9 (2020), 56-64.
5. Campion, M.A., Fink, A.A., Ruggeberg, B.J., Carr, L., Phillips, G.M., and Odman, R.B. Doing competencies well: Best practices in competency modeling. *Personnel Psychology* 64, 1 (2011), 225-262.
6. Cateté, V. and Barnes, T. Application of the Delphi method in computer science principles rubric creation. In *Proceedings of the 2017 ACM Conf. on Innovation and Technology in Computer Science Education*, 164-169.
7. Connolly, R. Why computing belongs within the social sciences. *Communications of the ACM* 63, 8 (2020), 54-59.
8. Dalkey, N. and Helmer, O. An experimental application of the Delphi method to the use of experts. *Management Science* 9, 3 (1963), 458-467.
9. Eshet, Y. Digital literacy: A conceptual framework for survival skills in the digital era. *J. of Educational Multimedia and Hypermedia* 13, 1 (2004), 93-107.
10. Ferguson, R., Barzilai, S., Ben-Zvi, D., Chinn, C.A., Herodotou, C., Hod, Y., Kali, Y., Kukukska-Hulme, A., Kupermintz, H., McAndrews, P., Rienties, B., Sagy, O., Scanlon, E., Sharples, M., Weller, M., and Whitelock, D. *Innovating pedagogy*. The Open University (2017). <https://www.learnlab.org/p/182004>
11. Flowers, A. Data scientist: A hot job that pays well. Indeed Hiring Lab (Jan. 17, 2019). <https://www.indeedlab.org/2019/01/17/data-scientist-job-outlook>.
12. Furht, B. and Villanustre, F. Introduction to big data. In *Big Data Technologies and Applications*, Springer, Cham. (2016), 3-11.
13. Halevi, G. and Moed, H. The evolution of big data as a research and scientific topic: Overview of the literature. *Research Trends* 30, 1 (2012), 3-6.
14. Hasson, F., Keeney, S., and McKenna, H. Research guidelines for the Delphi survey technique. *J. of Advanced Nursing* 32, 4 (2000), 1008-1015.
15. Hastings, J.S., Howison, M., Lawless, T., Ucles, J., and White, P. Unlocking data to improve public policy. *Communications of the ACM* 62, 10 (2019), 48-53.
16. Hsu, C.C. and Sandford, B.A. The Delphi technique: Making sense of consensus. *Practical Assessment, Research, and Evaluation* 12, 10 (2007), 1-8.
17. Jo, E.S. and Geburu, T. Lessons from archives: Strategies for collecting sociocultural data in machine learning. In *Proceedings of the 2020 Conf. on Fairness, Accountability, and Transparency*, 306-316.
18. Jones, W., Capra, R., Diekema, A., Teevan, J., Pérez-Quinones, M., Dinneen, J.D., and Hemminger, B. 'For telling' the present: Using the Delphi method to understand personal information management practices. In *Proceedings of the 33rd Annual Conf. on Human Factors in Computing Systems* (2015), 3513-3522.
19. Keeney, S., Hasson, F., and McKenna, H. Consulting the

- oracle: Ten lessons from using the Delphi technique in nursing research. *J. of Advanced Nursing* 53, 2 (2006), 205-212.
20. Key findings from Gartner Marketing Analytics Survey 2018. Gartner (May 16, 2018). <https://www.gartner.com/en/marketing/insights/articles/key-findings-from-gartner-marketing-analytics-survey-2018>.
21. Lucia, A.D. and Lepsinger, R. *The Art and Science of Competency Models*, Jossey-Bass/Pfeiffer. San Francisco, CA (1999)
22. Marr, B. The 6 top data jobs in 2018. *Forbes Magazine* (2018). <https://www.forbes.com/sites/bernardmarr/2018/05/09/the-6-top-data-jobs-in-2018/#2ce81210430d>.
23. Martinez, W. How science and technology developments impact employment and education. In *Proceedings of the National Academy of Sciences* 115, 50 (2018), 12624-12629.
24. McAfee, A., Brynjolfsson, E., Davenport, T.H., Patil, D.J., and Barton, D. Big data: The management revolution. *Harvard Business Review* 90, 10 (2012), 60-68.
25. Metcalf, J., Keller, E.F., and boyd, d. Perspectives on big data, ethics, and society. *Council for Big Data, Ethics, and Society* (2020). <https://bdes.datasociety.net/council-output/perspectives-on-big-data-ethics-and-society/>
26. NIST big data public interoperability framework: Volume 1, definitions. National Institute of Standards and Technology (2015).
27. *The definition and selection of key competencies: Executive summary*. Organization for Economic Co-operation and Development; <http://www.oecd.org/pisa/35070367.pdf>.
28. Pellegrino, J.W. and Hilton, M.L. Education for life and work: Developing transferable knowledge and skills in the 21st century. National Research Council Division on Behavioral and Social Sciences and Education. Washington, D.C. (2012).
29. Radovilsky, Z., Hegde, V., Acharya, A., and Uma, U. Skills requirements of business data analytics and data science jobs: A comparative analysis. *J. of Supply Chain and Operations Management* 16, 1 (2018), 82.
30. Russom, P. TDWI best practices report: Big data analytics. TDWI (2011).
31. Shearer, C. The CRISP-DM model: The new blueprint for data mining. *J. of Data Warehousing* 5, 4 (2000), 13-22.
32. The O*NET-SOC taxonomy. O*NET Research Center (2019). <https://www.onetcenter.org/resource.html?c=Taxonomy>
33. Tissenbaum, M. and Ottenbreit-Leftwich, A. A vision of K-12 computer science education for 2030. *Communications of the ACM* 63, 5 (2020), 42-44.
34. Topi, H., Valacich, J.S., Wright, R.T., Kaiser, K., Nunamaker Jr., J.F., Sipior, J.C., and de Vreede, G.J. IS 2010: Curriculum guidelines for undergraduate degree programs in information systems. *Communications of the Association for Information Systems* 26, 1 (2010), 18. <https://eduglopedia.org/ais-global-is-education-report-2018-syllabus>.
35. Turoff, M. (1972). Delphi and its potential impact on information systems. In *Proceedings of the 1971 Fall Joint Computer Conference*, 317-326.
36. Watson, H.J. Tutorial: Big data analytics: Concepts, technologies, and applications. *Communications of the Association for Information Systems* 34, 1 (2014), 65.
37. Zook, M., Barocas, S., Boyd, D., Crawford, K., Keller, E., Gangadharan, S.P., Goodman, A., Hollander, R., Koenig, B.A., Metcalf, J., Narayanan, A., Nelson, A., and Pasquale, F. Ten simple rules for responsible big data research. *PLoS Computational Biology* 13, 3 (2017).

Orli Weiser (orli.weiser@gmail.com) is a Ph.D. candidate at the Industrial Engineering and Management Department at Ben-Gurion University of the Negev, Israel.

Yoram M. Kalman is an associate professor at the Department of Management and Economics, The Open University of Israel.

Carmel Kent is the head of educational data science at EDUCATE Ventures, U.K.

Gilad Ravid is an associate professor at the Industrial Engineering and Management Department at Ben-Gurion University of the Negev, Israel.

Copyright held by author(s)/owner(s).
Publication rights licensed to ACM.