Deloitte.
Insights

SIGNALS FOR STRATEGISTS

Democratizing data science to bridge the talent gap

David Schatsky, Rameeta Chauhan, and Craig Muraskin

Without data science, companies can't get full value from data, and there aren't enough data scientists to go around. But automation and training are giving companies access to data science without having to wage a war for talent.

AS COGNITIVE AND IoT technologies generate ever larger and more varied data sets, companies face the challenge of unlocking its value. And those that are failing to effectively apply data science may be putting themselves at a competitive disadvantage. *Data scientist* is one of the hottest job titles today, and battles for talent are fierce. Whether the talent shortage is real or over-hyped, companies should investigate a mix of new tools, staffing models, and training strategies.

Signals

- More than 40 percent of data science tasks are expected to be automated by 2020.[1]
- Early adopters of data science automation tools across industries are reporting significant time[2] and cost savings[3] as well as revenue gains.[4]
- Major technology vendors have introduced multiple tools to significantly simplify the application of data science techniques.[5]
- The market for low-code development platforms, which make application development and basic data science functions available to noncoders, is currently worth about US$4 billion globally.[6]
- Numerous training courses and "boot camps" have been launched to help professionals with basic mathematics and coding backgrounds acquire data science skills in days or months.[7]

The hottest job and the scarcest talent

The title *data scientist* generally refers to a professional with a graduate degree in computer science and expertise in mathematics, statistics, computer programming, and business knowledge.[8] These specialists tend to handle a variety of tasks critical to enterprise analytics projects such as collecting, cleansing, and organizing large and varied data sets; designing and testing various algorithms; building and deploying machine learning-based solutions; analyzing data for patterns; and communicating findings to business stakeholders.

Since nearly every major company is actively looking for data science talent, the demand has rapidly outpaced the supply of people with required skills.[9] (Based on current demand and supply dynamics, the United States alone is projected to face a shortfall of some 250,000 data scientists by 2024.[10]) Data science and analytics jobs typically take 45 days to fill, five days longer than the US market average, according to one study.[11] The skills gap and longer hiring times can cause project delays and higher costs, hindering enterprises' data analytics efforts. But a number of recent trends may change how companies acquire and apply data science capabilities, presenting savvy companies with some options for alleviating the talent bottleneck.

Automation and education are bringing data science to the masses

Most vendors in the data science and analytics market have made tool simplification a top goal; they are aiming to broaden and accelerate the adoption of data science and analytics capabilities. And an array of training resources is helping professionals with diverse backgrounds gain relevant data science skills. For the foreseeable future, elite data scientists will be in high demand. But five factors are beginning to democratize data science, helping to put this critical capability in the hands of more

professionals and potentially alleviating a crippling talent shortage.

Automated machine learning. By some estimates, data scientists spend around 80 percent of their time on repetitive and tedious tasks that can be fully or partially automated.[12] These tasks might include data preparation, feature engineering and selection, and algorithm selection and evaluation. Various tools and techniques designed to automate such tasks have been introduced by both established vendors[13] and startups.[14] Automating the work of data scientists helps make them more productive and more effective. Organizations can make aggressive use of data science automation to empower and leverage oversubscribed talent.

App development without coding. Low-code and no-code software development platforms offer graphical user interfaces, drag-and-drop modules, and other user-friendly structures to help both IT and nontechnical staff accelerate AI app development and delivery. For example, using a no-code platform, salespeople can build a machine learning-based tool themselves to provide product recommendations to customers based on cross-sell opportunities. These platforms have the potential to make software development up to 10 times faster than traditional methods.[15] Apart from building their own solutions, key technology vendors[16] have acquired startups to offer or strengthen their low-code and no-code platforms. One analyst firm has estimated that the market for these platforms is growing at 50 percent annually.[17]

Pre-trained AI models. Developing and training machine learning modules is a core activity of data scientists. Now, key AI software vendors[18] as well as several startups[19] have launched pre-trained AI models, effectively packaging machine learning expertise and turning it into products. These solutions can slash the time and effort required for training,[20] or even start producing specific insights right away.[21] Mostly pre-trained models are available for use cases related to image, video, audio, or text analysis such as sentiment analysis,[22] sales opportunity workflow automation,[23] customer service,[24] automated equipment inspection,[25] and

interactive advertising.[26] We can expect more pre-trained models to come to market in coming months.

Self-service data analytics. Increasingly, business or nontechnical users have tools at their disposal that can deliver data-based insights without involving analytics specialists, including data scientists. Self-service analytics tools offered by many business intelligence and analytics vendors[27] now include features to augment data analytics and discovery. Some automate the process of developing and deploying machine learning models. Features such as natural language query and search, visual data discovery, and natural language generation help users automatically find, visualize, and narrate data findings like correlations, exceptions, clusters, links, and predictions. These capabilities empower business users to perform complex data analysis and get quick access to customized insights without relying on data scientists and analytics teams.

Accelerated learning. Data science and AI-related training courses and boot camps are proliferating.[28] These training programs are aimed at professionals with basic mathematics and coding backgrounds and can impart basic data science skills in a period ranging from a couple of days to a couple of months. Such courses are intended to enable professionals to bring basic data science skills to projects quickly.

CHANGING ROLES OF DATA ANALYTICS AND BUSINESS TEAMS

Many organizations don't recognize the mix of talent and skills required to be successful when applying data science. Some put great faith in data scientists but fail to reckon with the importance of business and functional expertise to the success of a project. A properly staffed initiative may include design-thinking skills to help conceptualize a solution, functional domain knowledge to help identify high-value use cases and shape the solution, business skills to articulate a compelling business case, data engineering skills to provide access to the right data in the form needed, and, for AI projects, AI skills to drive execution of a variety of AI technologies.

Success depends on more than technology talent—it requires the right mix of skills and expertise.[29]

Eventually, the democratization of data science will enable greater collaboration between business and data science experts in building data-centered solutions. Some companies have started effectively expanding their data science efforts by providing data science automation tools to a mix of professionals including data scientists, data engineers, statisticians, and business users.[30] Others find that breaking down the data science role into a collection of more specialized roles with overlapping skills makes it easier to get the mix of skills required to staff projects.

Challenges on the way to data science democratization

To benefit from the democratization of data science and analytics, enterprises need to first address certain challenges. Since a lot of the technological advancements in this area have happened recently, enterprises may encounter resistance to using these solutions. Business users may not be ready to trust them, preferring to continue relying on intuition and traditional decision-making processes. Technical experts, by contrast, may resist changing their workstyle and automating tasks they think of as requiring expert craftsmanship.

On the other hand, embracing the democratization of data science may present a different set of challenges. Without proper onboarding and training, users provided access to data science automation and self-service tools may fail to derive relevant insights or misinterpret or misapply the results in decision-making. Wide adoption of these tools will necessitate instituting governance procedures that run the risk of becoming bottlenecks. Inadequate data controls and governance practices in enterprises may lead to creation of information silos, bad analysis, and lack of accountability. Thus,

companies need to prepare to address these challenges before moving forward with data science democratization.

DEMOCRATIZING DATA SCIENCE IN YOUR ORGANIZATION

Companies seeking to develop data science capabilities are facing a tight market for talent. To avoid being blocked by a labor shortage, they should consider a multipronged approach, including employing automated tools and pre-trained models, empowering nontechnical users with no-code tools and self-service analytics, and investing in training their own staff in data science by selecting a high-quality, accelerated training option from among the many currently available.

Companies should also explore hybrid staffing models for their data science projects. Rather than overburdening the data scientists with all the analytics work, they can assemble combinations of experts such as data engineers, statisticians, and business analysts and equip them with relevant data science automation and self-service tools. Subject matter experts who can "speak data" to data scientists while "speaking business" to executives can be valuable additions to the teams working on data science projects.[31] This helps to foster a culture of collaboration between data science experts and business users, enabling data scientists to focus more on advanced and complex processes while reducing time to access actionable insights for business users.

Those enterprises that seek to build armies of data scientists may continue to struggle to hire the desired talent, end up overspending on salaries, and get stuck with excess human capital in coming years. Those that leverage new automation, self-service, and training solutions may be able to mitigate the data scientist shortage without going on a hiring binge.

Endnotes

1. Gartner, "Gartner says more than 40 percent of data science tasks will be automated by 2020," January 16, 2017.

2. Tas Bindi, "How machine learning is helping Virgin boost its frequent flyer business," ZD Net, November 21, 2017.

3. Heidi Miller, "Loyalty Lab implements machine learning without hiring a single AI data scientist," AWS Marketplace, March 7, 2018.

4. Salesforce, "Salesforce launches new low-code tools on the Lightning Platform empowering teams to collaborate and build apps fast," September 13, 2018.

5. Khari Johnson, "Microsoft introduces Azure service to automatically build AI models," VentureBeat, September 24, 2018; Conner Forrest, "IBM platform integrates data science and machine learning to boost enterprise AI," TechRepublic, March 16, 2018; IBM, "IBM largest ever AI toolset release is tailor made for 9 industries and professions," September 24, 2018; Kyle Wiggers, "Salesforce open-sources TransmogrifAI, the machine learning library that powers Einstein," VentureBeat, August 16, 2018; Larry Dignan, "Salesforce aims to bolster analytics for business users via natural language queries, easier visualization tools," ZD Net, March 5, 2018; Kyle Wiggers, "Google announces AutoML Vision, natural language translation, and contact center AI," VentureBeat, July 24, 2018.

6. John Rymer, "Siemens snaps up Mendix; low-code platforms enter new phase," Forrester, August 2, 2018.

7. *Economist*, "New schemes teach the masses to build AI," October 25, 2018.

8. Michael Sasso, "This is America's hottest job," Bloomberg, May 18, 2018.

9. Louis Columbus, "IBM predicts demand for data scientists will soar 28% by 2020," *Forbes*, May 13, 2017; Paul Petrone, "Why it's really good to be a data scientist right now," LinkedIn Learning Blog, August 20, 2018.

10. McKinsey Global Institute, *The age of analytics: Competing in a data-driven world*, December 2016.

11. IBM, "Quant crunch: How the demand for data science skills is disrupting the job market," December 29, 2017.

12. Armand Ruiz Gabernet and Jay Limburn, "Breaking the 80/20 rule: How data catalogs transform data scientists' productivity," IBM Cloud Blog, August 23, 2017.

13. Wiggers, "Google announces AutoML Vision, natural language translation, and contact center AI"; Forrest, "IBM platform integrates data science and machine learning to boost enterprise AI"; Johnson, "Microsoft introduces Azure service to automatically build AI models."

14. DataRobot website, accessed November 28, 2018; BusinessWire, "Feedzai unveils AutoML:Automated machine learning that fights fraud in a fraction of the time," August 7, 2018; Bonsai website, accessed November 28, 2018; Razorthink, "Razorthink delivers first deep learning data science automation platform," Globe Newswire, October 16, 2017.

15. John Rymer, "Why you need to know about low-code, even if you're not responsible for software delivery," Forrester, August 8, 2018.

16. Frederic Lardinois, "Microsoft acquires Lobe, a drag-and-drop AI tool," TechCrunch, September 13, 2018; Blair Hanley Frank, "Salesforce to acquire MuleSoft for $6.5 billion," VentureBeat, March 20, 2018; Frederic Lardinois, "Siemens acquires low-code platform Mendix for $700M," TechCrunch, August 1, 2018.

17. Rymer, "Siemens snaps up Mendix; low-code platforms enter new phase."

18. IBM, "IBM largest ever AI toolset release is tailor made for 9 industries and professions"; Google Cloud, "Cloud AI products"; AWS, "Machine learning on AWS"; Heidi Steen et al., "Pre-trained machine learning models for sentiment analysis and image detection," Microsoft, February 16, 2018.

19. Indico, "Indico IPA"; Nanonets, "Drone model demos"; Hive.ai, "Hive's deep learning platform supports companies through all stages of the machine learning workflow"; TwentyBN, "Pre-trained deep learning models," all accessed November 28, 2018.

20. Indico, "Indico IPA."

21. IBM, "IBM largest ever AI toolset release is tailor made for 9 industries and professions."

22. Steen et al., "Pre-trained machine learning models for sentiment analysis and image detection."

23. Indico, "Sales & support," accessed November 28, 2018.

24. IBM, "IBM largest ever AI toolset release is tailor made for 9 industries and professions."

25. Nanonets, "Drone model demos."

26. TwentyBN, "Pre-trained deep learning models."

27. Dignan, "Salesforce aims to bolster analytics for business users via natural language queries, easier visualization tools;" Tableau, "Tableau vs. traditional BI," accessed November 28, 2018; Jeff Martin, "Enabling self-service analytics with IBM Cognos Analytics and IBM Information Governance Catalog," IBM Community, April 8, 2017; Kyle Wiggers, "Adobe Analytics can now automatically surface insights for you," VentureBeat, September 24, 2018; Pam Baker, "Microsoft Power BI," *PCMag*, July 13, 2018.

28. Udacity, "Deep learning by Google"; Microsoft EdX, "Microsoft professional program in artificial intelligence"; Fast. ai, "Practical deep learning for coders, part 1"; Deeplearning.ai, "Break into AI," all accessed November 28, 2018.

29. For a study of executives' attitudes, behaviors, and investments related to artificial intelligence, see Jeff Loucks, Tom Davenport, and David Schatsky, *State of AI in the enterprise, 2nd edition*, Deloitte Insights, October 22, 2018.

30. Alex Woodie, "How to build a data science team now," Datanami, August 6, 2018.

31. Loucks, Davenport, and Schatsky, *State of AI in the enterprise, 2nd edition*.

About the authors

DAVID SCHATSKY is a managing director at Deloitte LLP. He tracks and analyzes emerging technology and business trends, including the growing impact of cognitive technologies, for the firm's leaders and its clients. Schatsky is on LinkedIn at www.linkedin.com/in/dschatsky and on Twitter @dschatsky.

RAMEETA CHAUHAN is a senior analyst at Deloitte Services India Pvt. Ltd. She tracks and analyzes emerging technology and business trends, with a primary focus on cognitive technologies, for Deloitte's leaders and its clients. Chauhan is on LinkedIn at www.linkedin.com/in/rameeta-chauhan-15b0311b.

CRAIG MURASKIN is managing director of the Deloitte US Innovation Group. He works with firm leadership to set the group's agenda and overall innovation strategy. Muraskin is on LinkedIn at www.linkedin.com/in/craig-muraskin-b7b2a629.

Acknowledgments

The authors would like to thank **Aniket Dongre** of Deloitte Support Services India Pvt Ltd., **Jonathan Camhi** of Deloitte LLP, and **Alok Ranjan** of Deloitte Services India Pvt Ltd.

Contact

David Schatsky
Managing director, US Innovation
Deloitte LLP
+1 646 582 5209
dschatsky@deloitte.com

Deloitte.
Insights

Sign up for Deloitte Insights updates at www.deloitte.com/insights.

Follow @DeloitteInsight

Deloitte Insights contributors
Editorial: Matthew Budman, Blythe Hurley, Abar Khan, and Rupesh Bhat
Creative: Molly Woodworth
Promotion: Nikita Garia
Cover artwork: Molly Woodworth

About Deloitte Insights

Deloitte Insights publishes original articles, reports and periodicals that provide insights for businesses, the public sector and NGOs. Our goal is to draw upon research and experience from throughout our professional services organization, and that of coauthors in academia and business, to advance the conversation on a broad spectrum of topics of interest to executives and government leaders.

Deloitte Insights is an imprint of Deloitte Development LLC.

About this publication

This publication contains general information only, and none of Deloitte Touche Tohmatsu Limited, its member firms, or its and their affiliates are, by means of this publication, rendering accounting, business, financial, investment, legal, tax, or other professional advice or services. This publication is not a substitute for such professional advice or services, nor should it be used as a basis for any decision or action that may affect your finances or your business. Before making any decision or taking any action that may affect your finances or your business, you should consult a qualified professional adviser.

None of Deloitte Touche Tohmatsu Limited, its member firms, or its and their respective affiliates shall be responsible for any loss whatsoever sustained by any person who relies on this publication.

About Deloitte

Deloitte refers to one or more of Deloitte Touche Tohmatsu Limited, a UK private company limited by guarantee ("DTTL"), its network of member firms, and their related entities. DTTL and each of its member firms are legally separate and independent entities. DTTL (also referred to as "Deloitte Global") does not provide services to clients. In the United States, Deloitte refers to one or more of the US member firms of DTTL, their related entities that operate using the "Deloitte" name in the United States and their respective affiliates. Certain services may not be available to attest clients under the rules and regulations of public accounting. Please see www.deloitte.com/about to learn more about our global network of member firms.