Data Science Team Topology: An Interdisciplinary Team Framework for AI

Ralitsa S. Maduro, PhD
Sentara Healthcare
Virginia Beach, VA
rali.maduro@gmail.com

Jay Gendron
Model The Cause
Chesapeake, VA
gerald.gendron@gmail.com

Phil Lacasse
Financial sector
San Antonio, TX
phil.a.lacasse@gmail.com

ABSTRACT

Artificial Intelligence (AI) technology has evolved, driven by continued harmony among scientific advancement, corporate interest, and human talent. This led many companies to adopt and develop AI solutions as part of corporate objectives to remain viable in the data age. Then, in the early 2020s, the US experienced the "Great Resignation," and employees across numerous industries left their jobs at levels not seen before. Workers in technology industries cited feeling professionally stagnant on teams that could not grow and working for managers who did not know how to support them. There is enthusiasm for using AI-based tools to alleviate some pressures of the Great Resignation by automating tasks as a solution for the tight labor market (Kuehner-Hebert, 2021), but such technologies require expertise to develop and maintain. Stabilizing the harmonious balance of AI growth, corporate adoption, and team development is as much a human resources issue as it is a technological one. With the accelerating pace of progress within the AI landscape, we propose that individual contributors and T-shaped employees (Guest, 1991; Schaffer, 2021) alone are insufficient to continue AI progress. Rather, by hiring T-shaped people and further developing them within a data science team topology, employers support more engaged employees who experience professional growth while providing the team capabilities needed to achieve corporate AI adoption objectives.

Keywords: T-shaped, team topology, AI ethics, interdisciplinary, Great Resignation, data science, causal inference

ABOUT THE AUTHORS

Ralitsa Maduro is a behavioral scientist who was working as an Enterprise Analytics Consultant for Research, Biostatistics, and Health Equity at Sentara Healthcare while she was preparing this manuscript. Prior to that, Dr. Maduro was a Biostatistician with the Quality Research Institute at Sentara Healthcare. Additionally, she is an Adjunct Faculty at the Department of Psychology at Virginia Wesleyan University. Dr. Maduro's new role is an Operations Research Analyst for the Navy Safety Center, Department of the Navy. Dr. Maduro holds a PhD in Applied Experimental Psychology from Old Dominion University, a Masters in Science Degree in Experimental Psychology from Old Dominion University, a Masters in Science in Clinical Psychology from Francis Marion University, and a Bachelors of Arts degree in Psychology from Stockton University.

Jay Gendron is a systems engineer who has worked as a data scientist in industries including software development, management consulting, national defense, and financial services. Jay served in the US Air Force for over 20 years as a program manager, educator, and operations research analyst. He holds a Bachelors of Science in Mechanical Engineering from the University of Maryland at College Park, a Masters of Science in Management of Technology from the National Technological University, and a Masters in Operational Science from the Air Force Institute of Technology. Jay is the author of the book *Introduction to R for Business Intelligence* and a number of chapters in edited titles. For more information, please visit www.linkedin.com/in/jaygendron.

Phil Lacasse is a Senior Data Scientist working for a major finance and insurance company that builds AI and machine learning solutions to serve its membership. He earned his Bachelors of Science in Chemical and Biomolecular Engineering from the Georgia Institute of Technology before taking a left-hand turn into data and began his data career gathering data sets in the oil and gas sector. He moved into an analytical role building statistical and physics-based models with an independent operator and during this time he earned his Masters of Science in Data and Analytics from the University of Texas at San Antonio before transitioning into the finance and insurance sector.

Data Science Team Topology: An Interdisciplinary Team Framework for AI

Ralitsa S. Maduro, PhD
Sentara Healthcare
Virginia Beach, VA
rali.maduro@gmail.com

Jay Gendron
Model The Cause
Chesapeake, VA
gerald.gendron@gmail.com

Phil Lacasse
Financial sector
San Antonio, TX
phil.a.lacasse@gmail.com

A NATURAL DISTRUST OF MODELS: PROJECTED DEMANDS OF ARTIFICIAL INTELLEGENCE

The need to maintain trust between the human user and artificial intelligence (AI) has been well documented (Jacovi & Goldberg, 2020; Lipton, 2018; Miller, 2019). Moreover, it has been argued that the only way to implement AI in our society safely is by designing **trustworthy** AI products (Jacovi, Marasović, Miller & Goldberg, 2021). Trustworthy may be defined using data science terms of reference: interpretability (the ability to understand the mathematical composition of a model directly), explainability (the ability to understand model functionality at the output level), and AI ethics (the ability to understand the bias that can result in discrimination and prejudice; Gendron & Maduro, 2022). Trustworthiness is becoming increasingly difficult to achieve due to the exponential increase in complexity and application of AI. Indeed, deciphering the "black box" has become nearly impossible while, at the same time, explainability has become more important (Castelvecchi, 2016). Data scientists often build solutions that fail one or more trustworthiness litmus tests. As Rai (2020) argues, explainable AI approaches should turn black-box models into glass-box models.

Mindfully building explainable glass-box models is essential in the context of existing cultural and socio-economic biases, nationally and globally. For example, Fletcher, Nakeshimana, and Olubeko (2021) argue that AI and machine learning (ML) must always be evaluated based on appropriateness, bias, and fairness when it comes to global health. Moreover, Suresh and Guttag (2021) formalized the framework for understanding sources of harm due to bias in AI and ML (Figure 1) and visualized the idea of baseline bias. For example, evidence suggests a race-based representation bias of testing for germline mutations in breast cancer that places black women at a disadvantage (McCarthy et al., 2016). Such disparity would negatively impact the accuracy of an AI system built on a model that uses genetic testing results as a predictor variable. In that case, the model will likely confound the risk of cancer due to the lack of race representativeness of the data.

There is increasing recognition that society needs to be protected against biased and unsafe AI. A United States National Institute of Standards and Technology initiative documented an increase in the legislative pressure on explainable models via a growing number of measures being introduced to study the impact of the use of AI (NCSL, 2022). Bills or resolutions written for AI were introduced in at least 17 states in 2021, with California having such a legislative proposal as early as 2019 (NCSL, 2022). Privacy concerns sparked similar regulations in the European Union as well. On April 21, 2021, the European Commission presented the Artificial Intelligence Act. As Kop (2021) states, "The Act seeks to codify the high standards of the EU trustworthy AI paradigm, which requires AI to be legally, ethically and technically robust while respecting democratic values, human rights, and the rule of law." A crucial part of the European AI Act is **trustworthy** AI by design. Elser (2022) notes that current legislative activities focus more on privacy protection (what cannot happen) rather than articulating what good faith AI includes. A critical part of developing future mindful AI models that lead to trustworthy AI is the interdisciplinary nature of the design teams (Kop, 2021). In this case, interdisciplinary is more than multidisciplinary – specifically in the connections across disciplines rather than mere knowledge of multiple domains (Rowland et al., 2018). Interdisciplinary teams are a prerequisite to trustworthy AI because no single employee can be an expert in all aspects of AI design.

There is a well-recognized lack of pedagogical content on domain knowledge and experience emphasis in training data science professionals (Bubb, 2020; Saha, 2020). Domain knowledge can help structure a model form leading to increased precision and accuracy of models. In the next section, we propose strategies for developing a data science

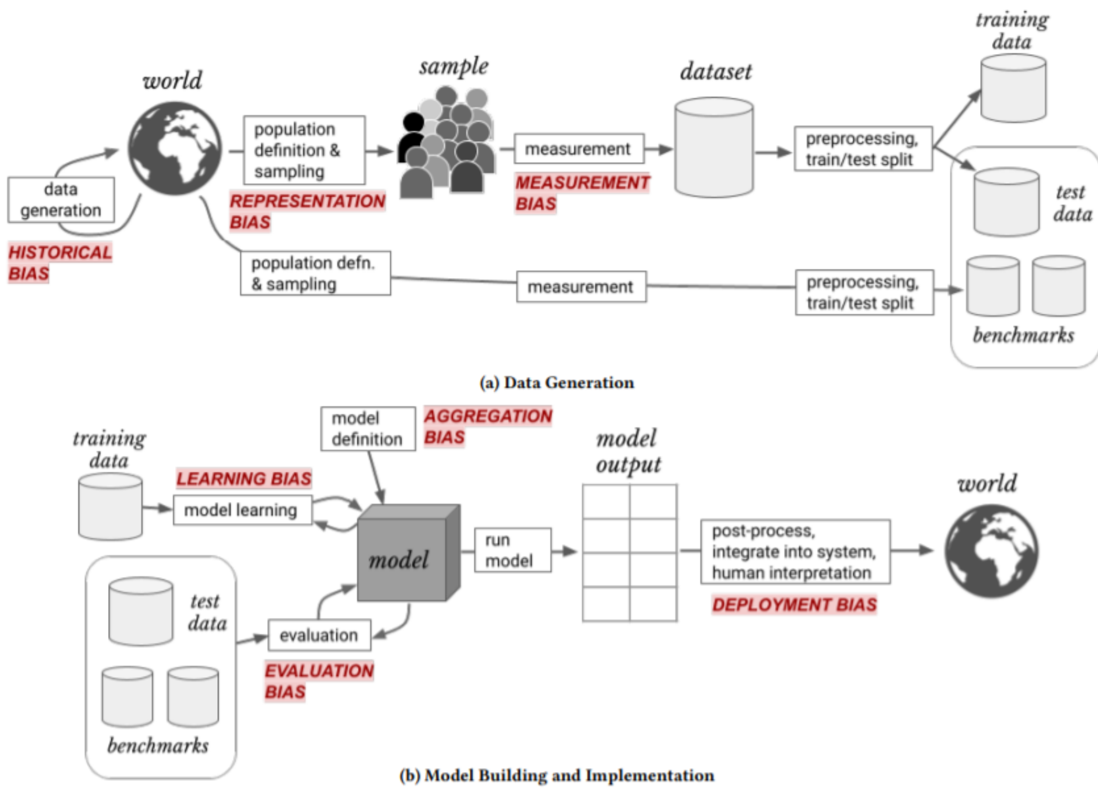

Figure 1: (a) The data generation process begins with data collection. This process involves defining a target population and sampling from it, as well as identifying and measuring features and labels. This dataset is split into training and test sets. Data is also collected (perhaps by a different process) into benchmark datasets. (b) A model is defined, and optimized on the training data. Test and benchmark data is used to evaluate it, and the final model is then integrated into a real-world context. This process is naturally cyclic, and decisions influenced by models affect the state of the world that exists the next time data is collected or decisions are applied. In red, we indicate where in this pipeline different sources of downstream harm might arise.

team topology of T-shaped data science employees as a means to hinder the arrival of a future AI winter[1] caused by unbalanced team skill sets.

DATA SCIENCE TEAM TOPOLOGY

Topology refers to "the way the parts of something are arranged and related" (Oxford Advanced American Dictionary, n.d.). The term topology is commonly seen in the technical literature referring to "network topology," but it more broadly relates to an entity described according to its relationships among component elements. As for a **team topology**, we define it as the way organizational members of a unit are assigned and related, one to another, in terms of their technical skills. Skelton and Pais (2019) use the term team topologies in their book of the same name to guide the formation of Dev-Ops teams in the software industry. We extend that concept to the related but fundamentally different data science domain. This leads to our working definition of **data science team topology**

> *The manner in which members of a data science team are assigned and related, one to another, in terms of their technical skills.*

This definition is primarily a logical design characteristic indicating the skill arrangement within a team. An assignment may be deliberate (during recruitment) or represent an existing team composition. Meanwhile,

[1] This AI winter will likely be the result of tightening resource availability as people and organizations lose interest, or fail to see the relevance of AI to their business challenges because the AI solutions were not trusted.

relationships can be described in many ways (e.g., personality type, communication styles, or degrees of interaction). In the context of a data science team topology, the relation of interest is the identification of technical skills possessed by team members such as:

- the individual
- a technical skill possessed by the individual
- the level at which that skill is utilized by that person within the team [primary subject matter expert (SME), ancillary SME, or general awareness]

Overall, the data science team topology captures the team's capabilities in relation to each other, not just as a listing of skills within a team.

Thinking in terms of the topology provides a richer sense of team capability, albeit with the cost of an added framework and associated analysis – a cost worth the outcome. Organizations generally recognize that talent drives AI success. Bubb (2020) notes talent as a leading risk to the enterprise, with over 80 percent of CEOs placing talent as their top corporate priority. Recall that we propose data science teams be deliberately formed and developed as topologies of T-shaped data science employees to continue the progress of AI into the coming decade.

Schaffer (2021) provides a useful chronology of the rise of T-shaped employees originating in the 1980s at McKinsey & Company. David Guest made this idea more relevant to engineering activities based on his work in 1991. By 2010, the CEO of IDEO, Tim Brown, described the "T-shaped person" in an interview while discussing hiring practices. Indeed, the T-shaped employee has been one of the well-worn organizing frameworks, whether as deliberate organizational design or endemic in hiring practices. Gardner and Spohrer (2020) note that the framework has also influenced the design of higher education. Figure 2 provides a representative visual of the T-shaped employee.

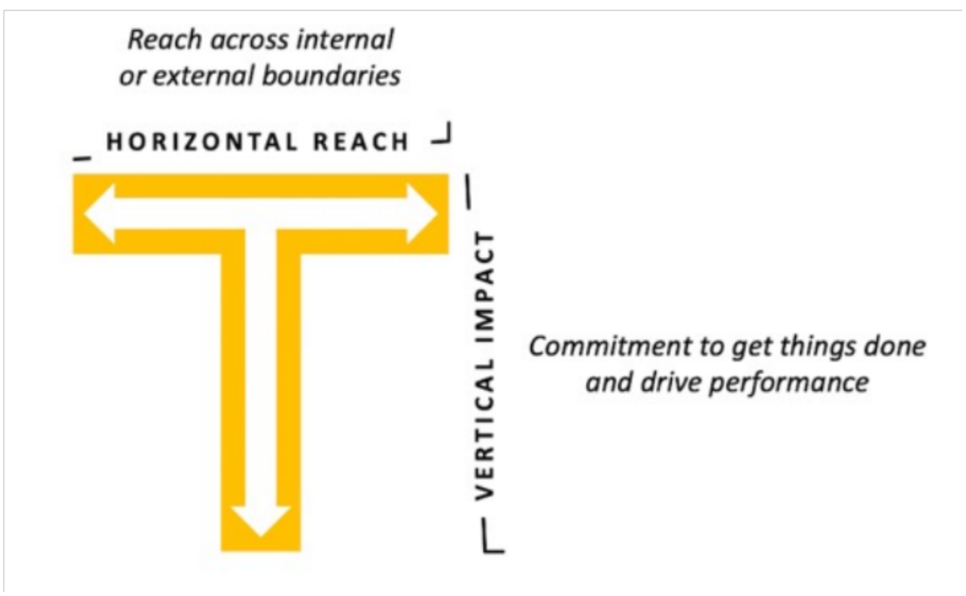

Figure 2. T-shaped employees have the perspective to manage across boundaries and I-shaped depth in skills to get things done (Schaffer, 2021).

Before the 2010s, there was little discussion of data scientists – although the sciences possessed many researchers proficient in the associated skills. Then came the infamous article by Davenport and Patil (2012) pronouncing the data scientist as The Sexiest Job of the 21st Century. By the mid-2010s, articles and blogs openly worried about the talent issues due to the lack of data science "unicorns" (individuals gifted with the skills now recognized as the hallmark of an entire team process). Evolution ensued, and the T-shaped data scientists emerged. Broadly speaking, a data scientist possesses math, statistics, and computer science skills with an appropriate awareness of domain knowledge, such as finance or medical (Conway, 2010). The data science team topology is built with such T-shaped data scientists.

A key benefit of data science team topologies based on T-shaped employees is the deeper systems thinking that results (Gardner & Maietta, 2020). A helpful analog to this train of thought is the domain of systems engineering. Systems engineering itself is an interdisciplinary approach to delivering value (Hutchinson et al., 2018). Conway (1968) presents the notion of a designed system – a sort of topology – that he expands into consideration of an organization as a system

> *Any system of consequence is structured from smaller subsystems which are interconnected. A description of a system, if it is to describe what goes on inside that system, must describe the system's connections to the outside world, and it must delineate each of the subsystems and how they are interconnected (p. 29).*

Eventually, we find that one of the most significant limiting factors to continued AI progress is how well AI practitioners (data science teams) organize into a value-producing system as perceived by AI advocates. Designing a data science team topology based on T-shaped employees as parts of a machine rather than components of a system suffers from sub-optimization issues unless you consider the system health as a whole. Stated another way: the data science recruiting process tends to optimize each hire individually (as expected) but does not complement another process to maximize overall team capability.

NEGOTIATING TEAMWORK BY DESIGN

While there is a great deal of focus on emerging technologies within the AI/ML landscapes, it is also critical for business leaders to remember that building, implementing, and maintaining those technologies requires a skilled and specialized labor force. The technology sector experiences some of the highest turnover rates of any industry, with estimates around 13% (Booz, 2018), and the direct and opportunity costs of executing a job search can add up quickly. Some sources estimate the total direct and opportunity costs of employee turnover in the organization to be between 100% and 300% of the position's salary (Ferrazi, 2015). Recruiting and retaining top technical talent had traditionally relied on the strength of a corporate brand: offering high salaries, excellent benefits, and luxury amenities (Larson, 2020). Increasingly, professional development and technological growth have been cited as key drivers for the millennial heavy tech sector workforce (Bean, 2019).

While many industries have clearly defined standards and examination schedules driving the associated career progression (e.g., taking the AICPA exam to become a certified public accountant), the technology sector has not established similar standards for data-driven roles. In an interview with *The Observer*, DJ Patil, often credited with popularizing the term "Data Scientist," describes the choice of title to be intentionally vague, citing that the critical thing to his perspective is "how you use data to interact with the world, study it and try to come up with new things." (Cao, 2019)

As such, managers and individual contributors often need to negotiate a technical development path that interests the practitioner while also serving the needs of the team and larger business strategy. This makes managing the team's technical topology a delicate and challenging task. If the negotiation fails, the business risks having gaps in their team's technical capabilities and even increased turnover as technical employees seek positions that will allow them to grow in their desired direction.

Considering the domain of systems engineering again, Project Helix began in the mid-2010s to catalog the critical skills of effective systems engineers. The project included organizations from the Department of Defense and the defense industrial base. It sought to model "the required skills of systems engineers outside of a T-model" but determined an alteration to the T-model was "an appropriate and useful framework" (Hutchinson, 2018). Figure 3 shows the Atlas Proficiency Model, a framework of skill dimensions, and a measurement system resulting from Project Helix.

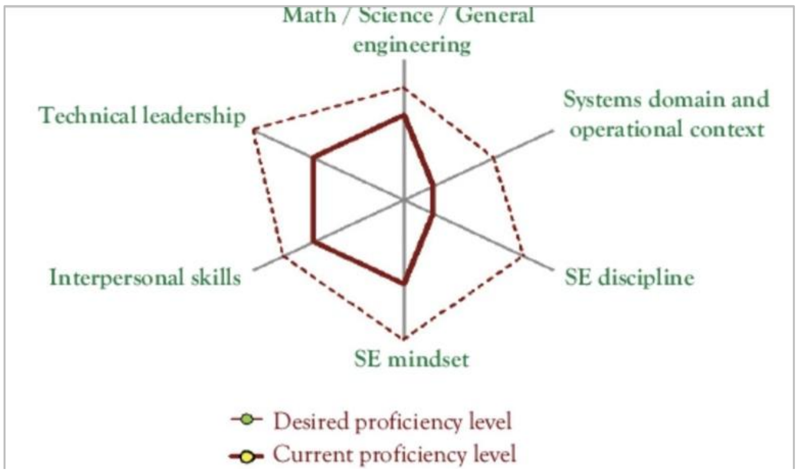

Figure 3. The Atlas Proficiency Model. Measurement system for an individual contributing systems engineer (Hutchinson et al., 2018).

Nonetheless, the Atlas Proficiency Model still captures a single technical proficiency (vertical of the "T") in "Math / Science / General engineering" axis as well as five softer skills (context, discipline, mindset, interpersonal, and leadership) that are not enablers to a topological approach. This is a helpful tool for measuring an individual (within the T-shaped framework), but we sought to consider measuring an **organization's situation** in its "team of T's."

This paper proposes a shift from the two-dimensional (breadth and depth) approach of the T-shaped employee model to a **coverage-based model**, focusing on the team's overall topology. To visualize this paradigm, we borrow the concept of ternary diagrams from the physical sciences. Popular in chemical engineering and geoscience, a ternary diagram is a two-dimensional representation of a mixture of three components. A common application of this analytical technique is describing the flammable mixtures of hydrogen, argon, and air in industrial processes (Nehrig et al., 2015). A flammability diagram is a standard tool used by chemical engineers in plant design because it allows the engineer to visualize which mixture regimes combustible-inert gas-oxidizer render the flammable substance effectively inert.

In Figure 4, we demonstrate how data science roles can be mapped into a combination of the three dimensions: Database Systems, Computer Science, and Statistics & Analytics. Within this framework, a practitioner's capability is represented by an area located by the three axes. These three dimensions are selected because they can represent a variety of AI/ML tasks in an industry-agnostic way. There are other dimensions of capability (such as design and visualization, industry-specific domain knowledge, soft skills, etc.) to also consider when mapping growth paths for practitioners.

In this view, formal education programs result in an expected, and consistent, region of skill (cyan, diagonal striped region). This area of skill expands through the practitioner's own experiences, interests, and study (cyan, dotted region), transforming what begins as a well-defined region into a more amoeba-shaped capability that is unique to the practitioner.

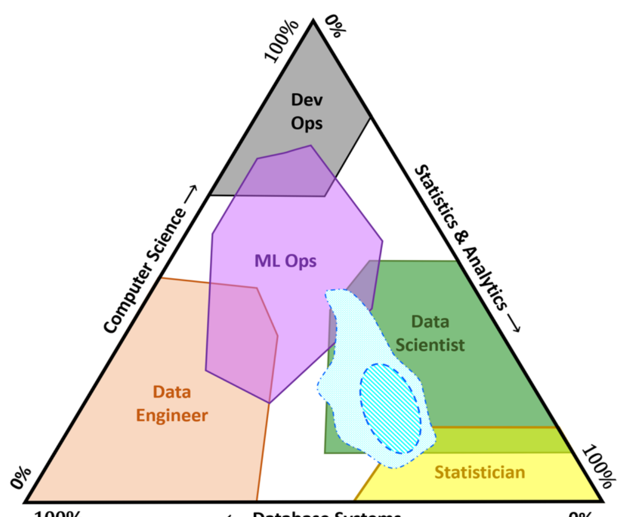

Figure 4. Framework. Ternary Diagram of Data Skills.

Many educational programs market their ability to create a "full-stack" individual who would be able to handle any task on the ternary diagram. However, this claim rings strongly of the "unicorn" approach to technology hiring. In practice, such practitioners are not trained; they are developed by a career's worth of experience.

Instead, when building or reshaping a team the key is to hire practitioners covering different regions of the framework. This is also the first opportunity to start practicing transparency in career and skill progression. Increasingly, hiring managers and candidates begin having frank discussions to level-set the needs and expectations of both sides of the business relationship. Companies like Paragus have decided to push the paradigm further by publishing "a clear view of [the practitioner's] personal career ladder in the form of a checklist of skills and experiences they need to earn before they are promoted" (Bean, 2019).

As the team coalesces and develops, the individual practitioners share knowledge and experience and their regions within the framework migrate towards one another. Observing this migration allows the team to anticipate one another’s needs and smooth out capability transitions. This familiarity reduces the time to market and improves the quality of the final products developed by the collaboration.

There is no one pattern of coverage that fits the needs of all organizations. Rather, the coverage should be determined by the organization's data maturity and goals. In the first stage of data maturity (**business reporting**) – organizations are often heavy in data engineers and business intelligence analysts. They seek to gather and communicate information and build the infrastructure and data assets needed for more advanced analysis and projects.

As the organization progresses into the **ad hoc analysis** level of maturity, there is a greater need for analysts and statisticians to discover and deliver insights. The organization may ultimately grow into **predictive and prescriptive analytics**, using ML engineers and data scientists to drive business planning and advanced ML products (Sisense, n.d.). ML-Ops and Dev-Ops practitioners can be very useful when the business directly sells ML products rather than using them for internal purposes. The maturation journey is unique to an enterprise. It does not require all roles to be filled. Neither does it require roles be filled in the same proportions at any given time in the company's data maturation process.

When gaps in the team's data science topology are identified, the first instinct may be to encourage the existing practitioners to grow in that direction. While economical, this approach assumes there is sufficient bandwidth for a period of learning and interest on the part of an existing team member to shift roles. If timing permits, another option would be to identify data-minded individuals within the organization that can be trained to fill the niche. Many companies have educational assistance programs that can assist in this growth. Assuming that appropriate and willing mentors can be found amongst existing practitioners, internal training may also be an option. Though data-minded individuals may be tricky to identify, they benefit from familiarity with the corporate culture and mission and more extensive domain knowledge than an external hire. Furthermore, their transition into a data role often represents career and salary growth, two commonly cited reasons for job dissatisfaction in the tech industry.

If an external hire must fill the role due to the specialized nature of the job or the timing requirements, then finding qualified candidates who have industry experience may be challenging. Many bootcamps and degree programs focus on data skills, assuming (often rightly) that domain familiarity will be obtained in the first few years of employment. Some candidates have the personality and motivation to seek out business SMEs to onboard that knowledge. However, others may benefit from a "business internship," a dedicated period of three to six months at the beginning of their time with the company to work within one or more of the lines of business. The authors have observed this approach to be particularly effective for early career hires.

Finally, not all capability gaps require a permanent hire. For example, a product team focused on building a natural language project for transcribing medical notes during a procedure may not benefit from having a full-time medical doctor on staff with the AI/ML team. Instead, engaging in consultation and partnership with a medical doctor(s) working for another part of the company would make sense. If the project progresses to a point where it deals with subspecialties of medicine, then contracting with external SMEs may be sufficient to provide the insight and guidance required. An added benefit of the approach is that engagement with SMEs early and often will help shape the final product into something more useful and intuitive for the planned end-user.

In addition to hiring and development, interdisciplinary data science team topologies benefit from a vision of where AI is moving in the next decade. Gendron and Maduro (2022) investigate the role of psychology in data science resulting in an approach termed mindful modeling based on the growing intradomain influence of causal inference.

SOCIAL SCIENCE, MINDFUL MODELING, AND CAUSAL INFERENCE IN AI

Mindful modeling promotes the use of a metamodel phase prior to data collection and computational modeling activities. The benefits of conceptual modeling are clearly described in prior work (Robinson et al., 2015) and point to the need for deliberate and structural thought processes that ultimately lead to causal inference. The importance of causal inference in AI could not be overstated, and as noted in AI Med (2021), "The future of AI, especially as we aim to reach artificial general intelligence, will need to learn similarly to how a child learns: understand casual relationships without overabundant data." Studies and work in causality emerged in the 1970s and were propelled into academics by Judea Pearl in the 1980s (Pearl et al., 2018). Unfortunately, causal inference methods remained largely unused by AI researchers and practitioners. We argue that AI teams should incorporate causal inference capabilities while striving to attain artificial general intelligence (Hünermund et al., 2019; Schölkopf et al., 2021). Together this creates the conditions to experience competitive advantages in data science organizations.

A type of modeling that leads to causal inference is structural equation modeling (SEM). Identifying latent variables that are believed to exist but cannot be directly observed (i.e., depression, decision-making, cognitive strain) is a burgeoning topic amongst social scientists and economists. Additionally, SEM type estimations are beneficial because they produce effect size estimations (such as R^2) that provide information about the magnitude or direction of results. Knowing how big or small the effect size is can inform practitioners about the relative, real-world value of a statistically significant, well-fitting model. Statistical significance on its own does not convey real-world value. For example, a hypothetical model suggesting that lab biomarkers can predict sepsis in patients may have excellent fit, yet little clinically significant value if the model only predicts 5% of the variance in sepsis.

Most social scientists and many behavioral economists often undergo extensive SEM training in their programs. However, AI researchers and practitioners are rarely trained in SEM. This mismatch between the goals of AI and the AI practitioner training points to the benefit of an intradisciplinary team in AI, particularly bringing social scientists to the table. Including SMEs from other disciplines such as psychology, mathematics, linguistics, philosophy, and experts in the field for which a product is developed (e.g., nurses and physicians for a healthcare AI), will benefit the AI team's product (Nahar et al., 2021). The importance of having the product users at the table from the onset is twofold: it increases trust and explainability and increases the likelihood of product users embracing AI. Lastly, promoting diversity, equity, and inclusion within each team and organization would further tear down silos, protect against implicit and explicit bias, and inspire innovation (Gendron & Maduro, 2022).

SUMMARY

By hiring T-shaped people and employing them in a thoughtfully designed data science team topology, employers can realize an environment where team members behave collectively to increase capability and appropriate skills. What can be done to avoid another AI winter? Turning black box models into glass box models is a hurdle that AI practitioners must overcome to achieve the usability and trustworthiness of their products. Regarding building teams that create glass box products, it is up to the organizational leaders to evolve and work towards a mature data science team. Designing a topology is more effective if managers can visualize or measure progress in its evolution. Development, hiring (internal and external), and consultation are three effective strategies to fill out a data science topology. Using a ternary diagram depicting coverage of three desired dimensions allows a glimpse into the overall coverage of team skill and capability relative to objectives. Mindful modeling techniques that include aspects of the scientific method, such as causal inference, will provide a competitive advantage to data science teams. We support the assertion that due to the general lack of formal education about causality in data science education, training will first be driven by AI practitioners using causal inference in corporate settings, which will inspire an evolution among coworkers (Hünermund et al., 2019; Hünermund, Copenhagen Business School, personal communication, January 2022). Lastly, we caution that leadership efforts towards advanced analytics, without the necessary underlying team infrastructure and data, are likely going to be ineffective.

REFERENCES

AI Med. (2021, October 26). *Causal inference: The future paradigm of artificial intelligence*. Feature: AC observations. https://ai-med.io/ac-observations/causal-inference-the-future-paradigm-of-artificial-intelligence/

Bean IV, D. D. (2019, December 11). *A small tech company tried it all to stop employee turnover. Only one thing worked.* CNBC. https://www.cnbc.com/2019/12/03/a-tech-firm-tried-it-all-to-stop-turnover-only-one-thing-worked.html

Booz, M. (2018, March 15). *These 3 industries have the highest talent turnover rates*. Linked-In. https://www.linkedin.com/business/talent/blog/talent-strategy/industries-with-the-highest-turnover-rates

Bubb, D. (2020). Forward. In P. Gardner & H. Maietta (Eds.), *Advancing talent development: Steps toward a T-model infused undergraduate education*. Business Expert Press.

Cao, S. (2019, November 9). *What on earth is a data scientist? The buzzword's inventor DJ Patil spills all.* Observer. https://observer.com/2019/11/data-scientist-inventor-dj-patil-interview-linkedin-job-market-trend/

Castelvecchi, D. (2016). Can we open the black box of AI? *Nature News, 538*(7623), 20.

Conway, D. (2010, September 30). *The data science Venn diagram*. http://drewconway.com/zia/2013/3/26/the-data-science-venn-diagram

Conway, M. E. (1968, April). How do committees invent? Design organization criteria. *Datamation,* 28-31.

Davenport, T. H., & Patil, D. J. (2012). Data scientist: The sexiest job of the 21st century. *Harvard Business Review, 90*(10), 70–76. http://hbr.org/2012/10/data-scientist-the-sexiest-job-of-the-21st-century/

Elser, J. (2022). *How to talk to data scientists: A guide for executives.* Business Expert Press.

Ferrazzi, K. (2015). Technology can save onboarding from itself. *Harvard Business Review*, *93*(3). https://hbr.org/2015/03/technology-can-save-onboarding-from-itself

Fletcher, R. R., Nakeshimana, A., & Olubeko, O. (2021). Addressing fairness, bias, and appropriate use of artificial intelligence and machine learning in global health. *Frontiers in Artificial Intelligence*, *3*, 116.

Gardner, P., & Maietta, H. (2020). Deep systems knowledge brings system thinking. In P. Gardner & H. Maietta (Eds.), A*dvancing talent development: Steps toward a T-model infused undergraduate education* (Ch. 2). Business Expert Press.

Gardner, P., & Spohrer, J. (2020). Introduction. In P. Gardner & H. Maietta (Eds.), *Advancing talent development: Steps toward a T-model infused undergraduate education*. Business Expert Press.

Gendron J., & Maduro, R. (2022). The role of inference in AI: Start S.M.A.L.L. with mindful models. In F. A. Batarseh & L. Freeman (Eds.) *AI assurance* (Ch. 6). London, UK: Elsevier Academic Press.

Guest, D. (1991, September 17). Managers in focus as the skills gap closes: The hunt is on for the renaissance man of computing. *The Independent*. London, England. p. 27, Issue 1536.

Hünermund, P., & Bareinboim, E. (2021, March 10). *Causal inference and data fusion in econometrics*. ArXiv Econometrics. arXiv:1912.09104v3 [econ.EM].

Hutchinson, N., Pyster, A., & Henry, D. (2018). From "T" to "∏": The other leg that systems engineers stand on. In Y. Moghaddam, H. Demirkan, & J. Spohrer (Eds.), *T-shaped professionals: Adaptive innovators.* Business Expert Press.

Jacovi A., & Goldberg, Y. (2020). Towards faithfully interpretable NLP systems: How should we define and evaluate faithfulness? *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 4198–4205. Association for Computational Linguistics. 10.18653/v1/2020.acl-main.386

Jacovi, A., Marasović, A., Miller, T., & Goldberg, & Y. (2021, March). Formalizing trust in artificial intelligence: Prerequisites, causes and goals of human trust in AI. *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 624-635. Association for Computing Machinery. https://doi.org/10.1145/3442188.3445923

Kop, M. (2021, October 1). *EU artificial intelligence act: The european approach to AI, transatlantic antitrust and IPR developments.* https://law.stanford.edu/publications/eu-artificial-intelligence-act-the-european-approach-to-ai/

Kuehner-Hebert, K. (2021, December 3). *How the 'Great Resignation' is accelerating the use of robotics and AI-based tools.* BenefitsPro. https://www.benefitspro.com/2021/12/03/how-the-great-resignation-is-accelerating-the-use-of-robotics-and-ai-based-tools/

Larson, B. (2020, May 4). *The biggest 2020 trend in the tech industry is the people that power it*. Lucas Group. https://www.lucasgroup.com/your-career-intel/biggest-2020-tech-trends/

Lipton, Z. S. (2018, October). The mythos of model interpretability. *Communications of the ACM, 61*(10), 36–43. https://doi.org/10.1145/3233231

McCarthy A. M., Bristol M., Domchek S. M., Groeneveld, P. W., Kim, Y., Motanyajudy, U. N., Shea, A., & Armstrong, K. (2016, August 1). Health care segregation, physician recommendation, and racial disparities in BRCA1/2 testing among women with breast cancer. *Journal of Clinical Oncology, 34*(22), 2610-2618. doi:10.1200/JCO.2015.66.0019

Miller, T. (2019). *Explanation in artificial intelligence: Insights from the social sciences*. ArXiv Artificial Intelligence. arXiv:1706.07269v3 [cs.AI].

Nahar, N., Zhou, S., Lewis, G., & Kästner, C. (2021, October 19). *More engineering, no silos: Rethinking processes and interfaces in collaboration between interdisciplinary teams for machine learning projects*. ArXiv Software Engineering. arXiv:2110.10234v1 [cs.SE].

National Conference of State Legislations (NCSL). (2022, January 5). *Legislation Related to Artificial Intelligence*. https://www.ncsl.org/research/telecommunications-and-information-technology/2020-legislation-related-to-artificial-intelligence.aspx

Nehrig, M., Wille, F., Feldkamp, M., & Rolle, A. (2015, May). Aspect of gas generation caused by residual water inside ILW packages. In *Proceedings of the 10th International Conference on Radioactive Materials Transport and Storage*, RAMTRANS 2015. https://www.researchgate.net/publication/278020995_aspects_of_gas_generation_caused_by_residual_water_inside_ilw_packages_aspects_of_gas_generation_caused_by_residual_water_inside_ilw_packages

Oxford Advanced American Dictionary. (n.d.). Topology. In *Oxford Learner's Dictionaries.* Retrieved January 7, 2022, from https://www.oxfordlearnersdictionaries.com/us/definition/american_english/topology

Pearl, J., & Mackenzie, D. (2018). *The book of why: The new science of cause and effect*. Basic books.

Rai, A. (2020). Explainable AI: From black box to glass box. *Journal of the Academy of Marketing Science*, *48*(1), 137-141.

Robinson, S., Arbez, G., Birta, L. G., Tolk, A., & Wagner, G. (2015, December). Conceptual modeling: definition, purpose and benefits. *2015 Winter Simulation Conference (WSC)*, 2812-2826.

Rowland, G., & Hamilton, J. (2018). Developing the connective tissue of the T. In Y. Moghaddam, H. Demirkan, & J. Spohrer (Eds.), *T-shaped professionals: Adaptive innovators.* Business Expert Press.

Saha, S. (2020, March 4). *Domain knowledge — The second most important skill to have as a data scientist*. https://towardsdatascience.com/the-second-most-important-skill-to-have-as-a-data-scientist-2311f9efa143

Schaffer, C. (2021, March 16). *"Both/And" leadership: Combining the benefits of I- and T-shaped leaders*. Harvard Business Publishing Corporate Learning. https://www.harvardbusiness.org/both-and-leadership-combining-the-benefits-of-i-and-t-shaped-leaders/

Schölkopf, B., Locatello, F., Bauer, S., Ke, N. R., Kalchbrenner, N., Goyal, A., & Bengio, Y. (2021). Toward causal representation learning. *Proceedings of the IEEE*, *109*(5), 612-634.

Sisense. (n.d.). *Data Maturity*. https://www.sisense.com/glossary/data-maturity/

Skelton, M., & Pais, M. (2019). *Team topologies: Organizing business and technology teams for fast flow.* Portland, OR: IT Revolution Press.

Suresh, H., & Guttag, J. (2021). A framework for understanding sources of harm throughout the machine learning life cycle. *Equity and Access in Algorithms, Mechanisms, and Optimization,* 1-9.