Kaleidoscopic Perspectives on Practicum-Based Data Science Education

**Shuang Frost¹ Aleksandrina Goeva² Javin Pombra³ William Seaton⁴
Sara Stoudt^{5,6} Ana Trisovic⁷ Chris Wang⁸ Catherine Zucker⁹**

¹Marshall School of Business, Lloyd Greif Center for Entrepreneurial Studies, University of Southern California, Los Angeles, California, United States of America,

²Macosko Lab, Stanley Center for Psychiatric Research, Broad Institute of Massachusetts Institute of Technology and Harvard, Cambridge, Massachusetts, United States of America,

³Department of Computer Science, Graduate School of Arts and Sciences, Harvard University, Cambridge, Massachusetts, United States of America,

⁴Institute for Applied Computational Science, Harvard John A. Paulson School of Engineering and Applied Sciences, Harvard University, Cambridge, Massachusetts, United States of America,

⁵Berkeley Institute for Data Science, University of California Berkeley, Berkeley, California, United States of America,

⁶Statistical & Data Sciences Program, Smith College, Northampton, Massachusetts, United States of America,

⁷Institute for Quantitative Social Science, Harvard University, Cambridge, Massachusetts, United States of America,

⁸Graduate School of Arts and Sciences, Harvard University, Cambridge, Massachusetts, United States of America,

⁹Institute for Theory and Computation, Harvard-Smithsonian Center for Astrophysics, Harvard College Observatory and Smithsonian Astrophysical Observatory, Cambridge, Massachusetts, United States of America

Published on: Feb 25, 2021

DOI: <https://doi.org/10.1162/99608f92.ba6336bc>

License: [Creative Commons Attribution 4.0 International License \(CC-BY 4.0\)](https://creativecommons.org/licenses/by/4.0/)

The acquisition of practical data science skills is at the forefront of our minds as Early Career Board members. We come from a variety of professional backgrounds and have learned practical skills in different ways, some of which were more effective than others. Through their description of Boston University's M.S. in Statistical Practice (MSSP) program, Kolaczyk et al. ("Statistics Practicum: Placing 'Practice' at the Center of Data Science Education," this issue) inspired us to consider our own experiences with practical learning and how well it worked for our individual career goals. The perspectives below summarize our discussions on the utility, attraction, and value of a practicum-based data science education. Moderated by Catherine Zucker and Will Seaton, each of us shared our experiences, debated the merits, and clarified our opinions through these conversations. Participants included:

Aleksandrina (Aleks) Goeva, Postdoctoral Fellow at the Broad Institute of MIT and Harvard

Shuang Frost, Postdoctoral Fellow at Lloyd Greif Center for Entrepreneurial Studies, University of Southern California

Javin Pombra, Harvard College Class of 2022, Computer Science and Applied Math

Will Seaton, Master's in Data Science, Harvard University, Institute for Applied Computational Science

Sara Stoudt, Lecturer in the Statistical and Data Sciences Program at Smith College

Ana Trisovic, Postdoctoral Fellow at the Institute for Quantitative Social Science at Harvard University

Chris Wang, Senior Investment Manager, SoftBank China Venture Capital

Catherine Zucker, Postdoctoral Fellow, Center for Astrophysics | Harvard & Smithsonian

Will Seaton (WS): Welcome! MSSP sounds like a challenging and rewarding program, but it is different from how most of us studied. How did you learn practical data science and statistics skills?

Sara Stoudt (SS): Beyond the foundations learned in formal classes, I have acquired most of my statistical and computing skills from a variety of projects and internships, learning apprentice-style from more experienced statisticians, domain scientists, and data scientists. As an undergraduate and graduate student, I spent a total of five summers at the National Institute of Standards and Technology, a government lab, where I worked on a variety of statistical problems in metrology. As an undergraduate, I also took part in a research class and independent study projects where I worked on problems seeded from various external partners. As a graduate student, I gained experience working for an industry startup, nonprofit, and newspaper in various summer internships, and through taking a statistical consulting course.

Since my own experience has been learning by doing, this proposed style of practicum learning, incorporating both statistical consulting and more industry-focused data science work, appeals to me in its efficiency. Since luck and serendipity played a role in forging my own real-world connections, I appreciated how the program's preexisting partners relieved the burden on the student to piece together their own training via extracurricular experiences.

Javin Pombra (PB): As a current undergraduate, I have learned data science from three primary sources: classes, student clubs, and internships. Through core classes, I learned foundational statistical skills in areas like probability theory, linear algebra, Bayesian machine learning, and foundations of machine learning theory. Through internship work in finance and at an entertainment company, I gained skills in cleaning data and deploying models. Finally, through my consulting work in a student club—Undergraduate Harvard.ai—I learned how to communicate data science topics to a variety of audiences and gained an end-to-end understanding of data science projects. In short, I believe practica would be most similar to the consulting work I did and thus help me gain communication skills. However, I believe practica would not be as conducive to learning some skills such as model deployment and mathematical foundations of data science. Mathematical foundations often only come from intensive classes that are primarily focused on theory. Further, the short time cycles of practica would likely prevent many aspects of deployment for machine learning models in the industry.

Ana Trisovic (AT): During my school years, I was an enthusiastic attendee of school competitions, so extracurricular and external resources early on had a significant role in my learning. Though my background is in computer science, much of my data science knowledge has been acquired through online resources and classes. That is mainly because there were not yet any data science courses aside from mathematics, probability, and statistics at my university when I was an undergraduate. The first online course that I took was “How to Process, Analyze and Visualize Data” by MIT OpenCourseware back in 2012. This course taught me the basics of Python, which underpins much of my postdoctoral work today. Later on, I was able to intern with the data science team at Microsoft Development Center and with the LHCb experiment at CERN, which provided real-world experiences on the challenges of working with data. Finally, knowing how thrilled and grateful I was for being able to access the MIT resources from the other side of the globe, I became an advocate and contributor of open educational resources.

Catherine Zucker (CZ): How important have practical skills vs. theoretical knowledge been in your own careers? When was one more useful than the other?

Aleks Goeva (AG): To me, theory and practice in data science are two sides of the same coin, both having something to offer the other. Gaining knowledge and experience in both takes perhaps twice as long, but their synergy constitutes more than the sum of its parts. As luck would have it, I was a teaching fellow during the first year and a half of the [M.S. in Statistical Practice](#) (MSSP) Program at Boston University. While at the time I did not yet have the perspective to fully appreciate how different this program was compared to a standard

M.S. program in statistics, I certainly sensed that I was a part of something paradigm-shifting. At that point, I was a Ph.D. student, and my research direction was about as theoretical as they come within statistics. I had already been a teaching fellow for a variety of traditional mathematics and statistics undergraduate and master's-level courses, so I was excited to partake in something new that had a distinctly different rhythm.

Being a teaching fellow for MSSP was path-changing to my scientific career. I vividly recall finding myself coming to life in a whole new way when I engaged with domain experts and their practical statistical problems through the MSSP consulting service. As a result, I deliberately decided to experience the applied side of statistics for my postdoctoral training and dove headfirst into biology. As the only statistician in my lab, I am glad to have a theoretical foundation while tackling the many exciting data-rich problems biology has to offer. Simultaneously, having spent time trying to answer real-world questions has given me a new appreciation and a deeper understanding of my theoretical work. Closely interfacing with domain experts and pairing models with data has also led me to make further theoretical contributions, such as discovering new equivalencies between dimensionality reduction techniques.

Chris Wang (CW): I think theory and practice are both important, but in different ways. When I was in business school, we had courses on data and analytics. The professors would receive data sets from companies and provide them to us in order to practice analytical methods. Even though I went through several iterations of this, some things remained unclear as I did not truly understand how the company worked and so could not come up with good research questions on my own. Working in the 'real world' as part of the venture capital industry, I have found that I need to do a lot of analytical work on the data that companies provide for due diligence. But first, I engage in deep conversations with the management team and industry experts to understand the context of the work. With this general background, I can better lead the data analytics work by proposing smart questions and hypotheses to dig in on. This higher level understanding helps a lot to improve the efficiency and effectiveness of the analytical work. The theory was necessary for me to discover new perspectives before engaging in the practice.

Shuang Frost (SF): My home discipline, anthropology, is a very academic and theoretical field. Most existing courses give students (both undergraduate and graduate students) an impression that academia is the only legitimate path for applying anthropological knowledge. I think practica are helpful because, first and foremost, they show students other contexts in which their knowledge and skills are valued. Second, practica place students in a diverse range of work environments (e.g., technology firms, traditional businesses, NGOs, thinktanks), enabling them to determine which career path best fits their goals. Lastly, practica could provide students with a better sense of which skills are needed for their desired industry (e.g., better quantitative skills or more effective communication). I see practical and theoretical lessons as complementary components of education. Theory gives students new ideas for innovation, a critical lens for understanding problems, and a holistic view of their work's influence on society, all of which are essential for tomorrow's youth. Therefore, practica could help them understand what real-world problems need solutions.

AG: How relevant is a practicum to your specific career path?

WS: I have spent the first part of my career in the technology industry, working with and advising both startups and large commercial institutions. An ever-present challenge is identifying talented people and pairing them with promising, relevant opportunities aligned with their skills. Namely, it is difficult to evaluate how well a recent graduate will fit into the company culture or whether someone claiming knowledge has had a chance to apply it. One of the best ways for employees to develop or prove valuable skills is by completing hard, short-term projects.

The authors' proposal for a practicum-based curriculum represents an excellent contribution for adapting project-based work to a graduate degree within an academic environment. It allows students to discuss tangible outcomes they have delivered, to understand the minutiae of daily challenges, and communicate this experience. Perhaps most importantly, it reveals to students what the 'first step' should be when actually trying to solve a real-world problem.

This ability to take the 'first step' is crucial if you are the lone data scientist within a team of subject-matter experts. You may have been hired for your statistical acumen, but must develop your own plan for how to bring data science techniques to bear on a practical application. You need to know what tools to apply and how to access data stored in different database types. You must be able to explain what you are doing to your teammates who might not have formal statistical training and how they can help you. Going through a practicum-based curriculum seems the best way to develop the ability to take that 'first step' and bring your teammates with you on the journey. The best way to do something in the future is to have done it before.

CZ: As an astronomer, practical training in data science is relevant for any future career in either academia or industry. Astronomy has entered the era of big data, with modern telescopes and numerical simulations of astronomical phenomena producing exabyte-scale data sets. To meet these data challenges requires an understanding of Bayesian analysis, machine learning techniques, database management, and large-scale computing. However, none of these skills are traditionally taught in undergraduate and graduate astronomy curricula. A data science practicum course would complement the traditional physics and math-based astronomy curricula, particularly if partnerships were made with industries that allowed merging data science skills with physics domain knowledge (e.g., aerospace or optics). Such a practicum would not only prepare astronomy Ph.D. students to tackle cutting-edge problems in modern astronomy research, but it would provide the necessary skills for a wider range of career paths.

WS: Each of us reacted differently when reading about the practicum so I am curious what your thoughts are on how to improve the program.

— **CZ:** Great question. For example, do you think short- or long-term projects are more conducive to learning?

JP: I can compare two experiences I had in my data science education. The first experience was Harvard's Advanced Practical Data Science Course (AC295), which taught concepts through two-week practica without external partners. The second experience was through data science internships, with a focus on long-term practical experience.

I found that long-term projects allow for more exploration, while short practica allow students to meet predefined goals. For instance, in AC295 practica the projects are fully mapped out and do not leave as much room for creativity. Short-term projects preclude computationally expensive techniques like some machine learning models that require long training times. Since the practica take place early in a student's academic career, as opposed to capstones, there is the possibility students have not acquired all the skills needed to successfully execute the project. In contrast, a long-term capstone in a student's final year may not cover as many data science skills as practica. Practica may also produce a larger technical portfolio for potential jobs.

SS responds: How Javin describes the depth versus breadth tradeoff resonates with me. It takes time to decide *what* to do and *why* for a project but we also need time to implement our ideas. Balancing both depth and breadth in a practicum is always challenging, so the question becomes what is the most efficient way to teach and learn both, given time constraints of an academic program. To Catherine's comment about the value of a domain-specific version of this practicum, this style could assume a baseline level of domain knowledge and build from there. Since students would begin with subject-matter expertise, the time-intensive component of identifying relevant questions and goals of the project could be bypassed entirely.

WS: Did you feel you were effectively taught communication skills?

Communicating the story and relevance behind the data is often as important to data science as the analysis itself. The authors discuss MSSP's early emphasis on improving their students' science communication abilities. Our group has had mixed success with formal communications training offered through data science programs and so discussed when and how we believe that training is most effective.

JP: In terms of teaching communication skills, working with external partners is especially beneficial for understanding how to communicate data science to a general audience. I have found that communicating technical assumptions is key for all audiences to understand the scope of models and limitations of results, as there is potential for misuse of data science without this communication. I have also found that creating conceptual diagrams is an important aspect of data science communication. While presenting to external audiences is key, I do think it is equally important to do internal presentations with data science professionals, to be able to communicate nuances and technicalities to future employers or colleagues. Ultimately, my communication skills came from projects that pushed me to use different media (reports, slide decks, talks) and iterating and improving my presentations for different audiences.

CW: My experience with data science has been closely linked with tools for both data analysis and data visualization—analytical tools like Microsoft Excel, Alteryx, and presentation tools like Microsoft PowerPoint and Tableau. In my current work, I need to present the findings from data analytics to senior team members, whose time is limited. This made me think about how I can convey ideas in more concise ways. I have adapted the ‘pyramid structure’ for every presentation, which presents the conclusion first and supports it with high-level data. If the listener wants to learn more about something, then you can review that aspect of the analysis with a more detailed explanation. This process of learning via ‘trial and error’ enabled me to communicate more succinctly and to use more visuals in my presentations.

AG: I have learned a significant portion of my communication skills while practicing statistics, and to some extent, while taking and teaching classes. Here is what I have distilled: When it comes to the statistical aspect of the project, be a teacher to your clients and collaborators. Your goal is to convey what is possible—the assumptions and the limitations of a given statistical approach. But when it comes to the problem they bring you, be their student. Be your most focused and most curious self, listen carefully and ask questions. Your goal is to understand all the relevant details of the problem and the domain, and then to abstract it into something clear, feasible, and falsifiable, with data.

SS responds: Aleks’s student/teacher metaphor is a good one. Putting my curiosity in action, I like to ‘go to the data’ and see the data-generating process in action. I have also found that exposure to a variety of statistical problems in context has helped build intuition. Should I be worried about this assumption being broken? Does the data look weird? Should this result trigger a red flag in my underlying analysis? These are the questions you can only answer once you have seen enough of them ‘in the wild.’ The statistical consulting portion of the practicum provides a real-world foundation for this.

CZ: Should we be incorporating ethical discussions into these practical lessons?

The integration of courses on ethics into traditional data science programs has been a large focus of recent years. The importance of engaging with these questions early continues to be proven by ongoing events in technology, governance, and accountability. But can ethics be taught well through practical lessons, or is it most helpful through abstract exercises?

AT: Yes, certainly. I think that responsible and ethical use of data is of utmost importance and should be incorporated in data science programs. For example, research data should be shared to increase the pace of research and reduce duplication of effort, and especially so if it was collected using public resources. However, security, ethics, or privacy risks of data use or sharing must be identified and assessed. It is essential to mention that these risks do not only lie in human or user data. There are other types of information that, if recklessly shared, can cause harm, such as habitats on endangered species that can be poached or locations of archeological dig sites that could be looted or destroyed. As this is a complex subject, Open Data Institute offers a free tool to help researchers identify and manage ethical issues in their data project called [the Data](#)

[Ethics Canvas](#) (thank you to Katherine Mika for sharing this resource). Another project that may alleviate this tension is [OpenDP](#), which aims to allow a privacy-protected analysis of sensitive data. Finally, there are other considerations, such as understanding data ownership and licenses, which are often not straightforward. For example, for an online resource to be reusable, it needs to have an assigned license that would permit reuse and specify its conditions, and it is not enough to be only ‘available online.’ Also, different licenses apply to different resources like data and code. Therefore, I think that data scientists in both academia and the industry need to have at least some basic understanding of and training in data ethics.

AG: Ethical decisions can surface in extremely practical considerations and so must be incorporated. Take, for example, the preparation of a visual for your data. As the saying goes, a picture is worth a thousand words — data visualization is a powerful communication medium utilized in any and all ‘printed’ data science storytelling formats, from academic to news articles, from textbooks to blog posts. The choices available to us when wrangling data into a plot (e.g., choosing to exclude a portion of the data by deeming them ‘outliers,’ using logarithmic vs. linear color gradient, [scaling the dot radius vs. area](#) to a quantitative variable) can have [a dramatic effect](#) on the reader’s visual comprehension of the signal contained in the data. Perhaps it can be justified to ‘play’ with all those settings to hone the story that the data support. But there is a stark difference in attitude between making adjustments that help a figure be more legible or color-blind friendly, and adjustments that intentionally bias the story a plot tells. The agenda of the figure-maker is not easily discernible to the consumer. And since it is unlikely that reviewers or consumers would be double-checking what the effect of different plotting parameters is to the overall perception of a plot, often it is entirely in the hands of the data analyst to be conscientious, making visualization manipulation an ethical issue that students in (and consumers of) any form of visual data science should be conscious of and trained in.

With increasing frequency in recent years, data scientists are among the key actors in global processes with wide-reaching impact, such as [applying AI in the Criminal Justice System](#), [using AI to improve medical care and biomedical research](#), [using AI for hiring](#), etc., and as such, are empowered to speak up about ethical issues that might impact a multitude of individuals. Those include, but are not limited to, missing or inadequate representation of some groups in the data used to train a model that will later be applied to populations containing individuals from the un- or underrepresented groups. Having access to the raw data provides the privilege to assess and notice potential biases in the data, but also this privilege comes with an ethical responsibility to raise concerns, or even refuse to engage with projects with potential ethical issues.

WS: We have all been blessed with excellent professors, teaching staff, and administrators in our programs. What staff do you think a novel program like this needs to succeed?

SS: Staffing this kind of a course sounds like its own form of the data scientist unicorn problem. What early-career researcher has a combination of industry experience, a breadth of statistics and data science skills, and ample teaching and team management experience? There is precedent for industry postdocs, where graduates of Ph.D. programs get to focus on research motivated by company problems or questions, and teaching

postdocs, where recent graduates hone their teaching skills rather than having a research-driven experience. What if we combined the two and created a 2-year postdoctoral position where the first year is spent working for an established industry partner and the second year is spent teaching the practicum? One can even imagine a setup where the summers remain free for independent research as well.

Working within the current incentive structure, where academic papers are the currency of career advancement, while pushing the boundary of what ‘counts’ toward career advancement, is a delicate balance. However, after completing this 2-year postdoctoral position, fellows could decide they want to go into industry with a year of practical experience (both technical and managerial) under their belt—not to mention an industry connection to guide them. Alternatively, fellows could also decide they want to remain in academia buoyed by potential papers that come from both industry problems and pedagogical contributions. Those fellows who stay in academia could spread this type of practicum to their future institutions.

AG responds: Sara’s idea is fantastic! My only addition is that such unicorn postdoctoral researchers would also make excellent future MSSP Professors.

CZ: At the end of the day, how would you define success for a practicum?

It is a challenge to adapt practicum-style learning to an academic environment. Grading rubrics cannot be uniformly defined for all students since each group has a different partner, goals, and deliverables. How then should success be measured for students and the program overall?

WS: While I believe practical, project-based work greatly accelerates a student’s learning, curriculum designers must be honest about who receives the most value: the students do, not necessarily the partners. We should consider what kinds of problems would be selected by external partners for completion by student data scientists. At least in industry and under good management, all problems that are existential to a company’s survival and growth should be staffed by the most talented full-time employees. Anything less puts the company at risk. This leaves less important or ‘moonshot’ ideas as potential candidates for student practica. This is not necessarily bad! Students want practical experience in a context that allows for mistakes and learning opportunities, something that cannot happen on critical projects.

The lion’s share of the value to partners thus comes in two ways: a) infrastructure or b) intangibles. Student data scientists can contribute meaningfully to the essential but less fun infrastructure portions of data science: data cleaning, documentation, pipeline construction, and so on. However, projects focused too much on less fun aspects lose attractiveness, especially when an entry-level data analyst could provide similar training with the added incentive of being paid to learn.

Building a practicum that promotes intangibles for partners and students can lead to the best long-term results. For partners, participation increases brand awareness, invites talented recruits to experience company culture and provides a channel to inform academia on what skills to teach the next generation. For students,

partnership reveals professional expectations of data scientists, allows a sampling of different work environments and creates the potential for longer-term career mentors—and maybe, a job offer! The authors describe well the time it takes to scope projects and the unique staff composition required to support each cohort continuously. However, this uniqueness is an opportunity, not a support debt. Partners should be evaluated on their long-term employment potential—is this a place where students would want to work? One concern with the authors’ existing structure is that the current 15 to 20 student team for one partner does not provide each student with sufficient ability to interact with the partner and increase their potential for postgraduation employment. Investing in these long-term relationships can provide significant recruiting value to external partners and justify the costs of attending graduate school with a high chance a student will find happy employment upon graduation—the ultimate measure of success for this program.

SF: I do think practica help their partner organizations in discovering new problems, forming new ideas, and innovating solutions. Even inexperienced students might bring fresh perspectives to old problems. Perhaps their deliverables have to be further polished to be applicable, but these new ideas themselves are valuable. However, I think students should have the option to replace practica with internships, especially for students who have a clear idea where they want to work in the future, and such options are not available in practica. But given the idiosyncratic nature of internships, the instructors and students have to agree beforehand on deliverables and milestones to achieve in the internship to make sure that the students are truly learning something.

Monetary compensation could be used in combination with more intangible rewards, such as access to internships and jobs. But in some cases, monetary rewards could help students to both evaluate their worth in the industry and support their education. I know universities that deliver these monetary rewards in tuition payment, which might help circumvent policy hurdles and benefit students directly.

JP: I agree with Shuang that internships provide real-world experience most similar to industry experience. I maintain healthy skepticism that students would have similar internship experiences in terms of educational value given the variety of possible roles. One benefit of a practicum versus an internship is that students may be exposed to applications of data science not seen before or outside their sphere of awareness. For instance, consider a student who plans to work at a large technology company but finds interest in a nonprofit. Regardless, I believe it is vital that students have input when choosing their projects. It also does not sit well with me that students do not get paid for work that benefits for-profit companies. It would be valuable to consider tuition assistance for students on financial aid or in work-study programs. I would prefer external partnerships with nonprofit organizations where synergies could occur for the common good.

WS: Having gone through your own academic program, what do you wish were in an ideal practicum-based curriculum?

SS: Despite the variety and difference in time scale of my practicum-like experiences, the common theme is to practice not only being a statistician but also being a collaborator. All of my previous projects required me to converse and iterate with others regularly rather than work in isolation to solve a problem. Being a good collaborator requires that vague skill we discussed previously: ‘communication.’ I completely agree with the authors that teaching communication skills is an important part of the curriculum and would argue that it is nontrivial to teach well. This is partially because the skill is a bit amorphous (what does ‘communication’ really mean in the context of statistics and data science?) and partially because few of us have actually received formal training in this skill in the first place.

Taking a step toward making ‘communication’ more concrete, these are some skills I consider vital: asking the right questions of domain scientist collaborators to lead us in the right analytical direction; translating what our findings mean back to them; making sure the broader public knows the ‘so what?’ of our work. Effective communication also requires writing well for multiple audiences, visualizing data and results accessibly, and focusing on the narrative of the work.

To help with developing that narrative, Deborah Nolan and I have thought a lot about teaching statistical writing and communicating with data more broadly. We have a [book](#) coming out later this year that aims to guide students and practitioners alike through the writing and broader communication process. In the meantime, check out our “[Reading to write](#)” [article](#) for a sense of what we are thinking. We still do not have all of the answers. As the authors said, assessment is a challenge. How do we measure the quality of, and improvement in, communication? This is something I continue to think about, so reach out if you have thoughts.

SF responds: Yes, Sara highlights an important issue, since criteria for evaluation could hugely impact the learning and working experience. I would advocate for guaranteeing that a small percentage of peer feedback is included to assess students’ collaborative working capability. Qualitative feedback could help put a student’s deliverables in context.

AT: My ideal practicum curriculum would highlight handling the whole data lifecycle (collection to preservation), effective communication of results through visualization, and efficiently working in a collaboration. Most courses today emphasize data analysis and machine learning approaches, whereas my ideal practicum curriculum would also contain lessons on ensuring the results are reproducible, replicable, and reusable. In computational research, reproducibility refers to obtaining consistent results using the same input data, code, and methods. It should be taught as one of the major principles of the scientific method. In addition, learning how to understand and communicate research results through visualizations is critical, as Aleks and others have highlighted previously. Plots need to be descriptive and intuitive enough that even a layperson could interpret them correctly. Finally, in contemporary work environments, we are urged to work efficiently and collaboratively, but we need specific training to achieve that. Going through a course like “[the Missing Semester of your CS Education](#)” (which I would recommend) is helpful in the long term.

AG: Data science is a fast-paced and ever-evolving field that demands that one intentionally continues their education after graduation in order to stay up-to-date. Therefore, a practicum program should introduce students to the myriad of options for staying abreast with new developments after they depart. For graduates of a data science master’s program, I would recommend staying engaged with the field by attending meetups and seminars (my favorite is [Models, Inference & Algorithms](#) at Boston’s Broad Institute) or enrolling in online courses. There is a lot of useful data science media out there that you can find by sieving through Twitter, blog posts, YouTube, and podcasts. Check out “[Theory and Practice](#)” and “[Talking Machines](#)” first. And of course, make sure you keep reading *Harvard Data Science Review*!

Disclosure Statement

Shuang Frost, Aleksandrina Goeva, Javin Pombra, William Seaton, Sara Stoudt, Ana Trisovic, Chris Wang, and Catherine Zucker have no financial or non-financial disclosures to share for this article.

©2021 Shuang Frost, Aleksandrina Goeva, Javin Pombra, William Seaton, Sara Stoudt, Ana Trisovic, Chris Wang, and Catherine Zucker. This article is licensed under a Creative Commons Attribution (CC BY 4.0) [International license](#), except where otherwise indicated with respect to particular material included in the article.