

O'REILLY®

2015 Edition

Women in Data

Cutting-Edge Practitioners and Their Views
on Critical Skills, Background, and Education



Cornelia Lévy-Benheton
& Shannon Cutt



Strata+ Hadoop

WORLD

Make Data Work
strataconf.com

Presented by O'Reilly and Cloudera,
Strata + Hadoop World is where
cutting-edge data science and new
business fundamentals intersect—
and merge.

- Learn business applications of data technologies
- Develop new skills through trainings and in-depth tutorials
- Connect with an international community of thousands who work with data

Women in Data

*Cutting-Edge Practitioners and Their
Views on Critical Skills, Background,
and Education*

Cornelia Lévy-Bencheton and Shannon Cutt

Beijing • Cambridge • Farnham • Köln • Sebastopol • Tokyo

O'REILLY®

Women in Data

by Cornelia Lévy-Bencheton and Shannon Cutt

Copyright © 2015 O'Reilly Media, Inc. All rights reserved.

Printed in the United States of America.

Published by O'Reilly Media, Inc., 1005 Gravenstein Highway North, Sebastopol, CA 95472.

O'Reilly books may be purchased for educational, business, or sales promotional use. Online editions are also available for most titles (<http://safaribooksonline.com>). For more information, contact our corporate/institutional sales department: 800-998-9938 or corporate@oreilly.com.

Editor: Shannon Cutt

Interior Designer: David Futato

Production Editor: Nicole Shelby

Cover Designer: Ellie Volckhausen

Copyeditor: Jasmine Kwityn

Illustrator: Rebecca Demarest

February 2015: First Edition

Revision History for the First Edition

2015-01-26: First Release

2015-04-10: Second Release

While the publisher and the author have used good faith efforts to ensure that the information and instructions contained in this work are accurate, the publisher and the author disclaim all responsibility for errors or omissions, including without limitation responsibility for damages resulting from the use of or reliance on this work. Use of the information and instructions contained in this work is at your own risk. If any code samples or other technology this work contains or describes is subject to open source licenses or the intellectual property rights of others, it is your responsibility to ensure that your use thereof complies with such licenses and/or rights.

978-1-491-92301-6

[LSI]

Table of Contents

Women in Data: Cutting-Edge Practitioners and Their Views on Critical Skills, Background, and Education.....	1
Introduction	1
Profiles of Cutting-Edge Practitioners	4

Women in Data: Cutting-Edge Practitioners and Their Views on Critical Skills, Background, and Education

Introduction

Women in data and technology are no longer outliers or anomalies; they are entering the mainstream and excelling where technical skills, advanced education, and no small amount of personal tenacity and brilliance are the minimum requirements. That said, women are still an underrepresented minority in the disciplines of science, technology, engineering, and math, known by the acronym STEM.

To investigate and understand how and why some women do extremely well, we interviewed 15 women in data to learn what got them to their current level of success, exactly what motivated them to get there, and their views about opportunities for women in tech. We were very keen on hearing their recommendations about what needs to get “fixed” to close the tech gender gap for others.

We think you will find the stories shared during these interviews both interesting and inspiring. They reveal insights that will widen the path for other women analysts, engineers, mathematicians, and data scientists. These insights include:

- An update on the expanding role of the contemporary data scientist
- New attitudes toward women in data among Millennials

- Benefits of the data and STEM fields as a career choice for women
- Much needed and increasingly sought after remedies for closing the gender gap

Wondering what's new? The gender gap in tech is not news, but here's what is: it's shrinking. The underrepresentation of women in tech has garnered tremendous attention and support of late to the point where the continued existence of the numbers disparity has fostered a nation-wide movement to bring more women into technical fields. Starting with the feeder pipeline of education (from kindergarten to university) and continuing through to diversity issues in and beyond the workplace, bridging the gender gap in STEM and tech is now a nation-wide crusade and a very hot topic.

The groundswell of attention comes from every possible sector: public and private companies, national and local governments, associations, educators, parents, teachers, scientific organizations, media publicity, and trade groups. Emphasis is on correcting a range of loss and leakage issues that occur at multiple points along the career continuum. Extending from the type of coursework offered in schools, factors that discourage women from selecting and staying with tech include cultural bias, behavioral psychology, and gender stereotypes. Now through increased publicity, there is a definite assault on the gender gap issue.

Our data practitioners confirm that dispelling myths of women's inability to do well in math and tech is only a small part of the battle. Other challenges center on advancing the idea that gender diversity fuels creativity, innovation, and economic growth. Much work needs to be done to publicize these truths and change the prevailing mindset.

Because women represent over 55% of the workforce, it is striking that fewer than 25% of jobs in technical and computing fields are held by women. When 58% of bachelor's degrees are being awarded to women, why are only 18% of computer science degrees going to women? Silicon Valley companies are leading the way in looking into these disparities and opening up advancement to better paying, higher prestige, leadership positions to their female employees. These jobs are also exciting and satisfying, and contribute handsomely to the bottom line.

Perhaps because big data has created a tsunami of new challenges and opportunities, or perhaps because of the well-publicized need to fill over 1.4 million new jobs in computer science by 2020 (jobs that will largely go unfilled), or perhaps because of a national sense of not wanting to fall behind on the world stage, closing the gender gap in tech is finally making it to the national priority list.

Our interviews with practitioners in data and STEM reveal that they are themselves the solution and model for the much needed changes that will help close the gender gap in tech.

Profiles of Cutting-Edge Practitioners

Carme Artigas

President/Co-Founder/Partner, Synergic Partners

MBA, Industrial Management & Administration, University of Ramon Llull

MS, Chemical Engineering, IQS - University of Ramon Llull



For Carme Artigas, a drive for innovation and an insistence on practical application are the art and science of professional life. With a diverse background ranging from cancer research in photochemistry and radiation at the Max Planck Institute, to investigating sustainable employment practices in Costa Rica, and running her own company in the field of big data and data science, Artigas reveals an impressive model for anyone seeking success in the business world, no matter the industry. When asked for a piece of advice for striving entrepreneurs, Artigas put it simply: “Convert your project, your work, into your life project. Don’t run a business because you want to be rich. Your work must be your passion.”

A passion for the transformational power of science and the draw to understand structure and function in the natural world are what led Artigas to pursue and complete a Masters of Science degree in Chemical Engineering, at the University of Ramon Llull in 1991. As a child, language and other soft sciences came very easy to Artigas—too easy, in fact. It was the challenge of studying science that appealed to her spirit of innovation. While growing-up in Spain, Artigas noted that there were not many prominent female role models in the sciences, and most young women chose public sector roles, in law, education, and medicine—these seemed to be a more natural expression of the social and cultural life at the time. In the Mediterranean culture in which she grew up, students were encouraged to study language, history, and other social studies; there was a

strong emphasis on being able to express your ideas, with much less focus on the importance of math and science. Cultural influence was no match, though, for Artigas' clear vision of science as a field capable of transforming the world; she knew that through science she could make an impact in any field—beginning with medicine.

After completing her Masters degree in Chemical Engineering, Artigas began a year of cancer research at the Max Planck Institute in Germany in 1991. There, she initiated a new line of research, studying the photosensitizing properties of cancer radiation treatment. Her goal was to determine how to use radiation to selectively treat cancerous cells, without destroying healthy cells. The results of her work were corroborated by other research teams, and were published in [an article](#) in the Elsevier *Journal of Photochemistry and Photobiology* in 1997. The article was later cited by four additional research projects related to cancer therapy. Artigas recollects this milestone as a true expression of her passion for science, and her unrelenting desire to make a difference in the world.

Faced with a critical decision in 1992—whether or not to continue her research and pursue a PhD in science, Artigas chose a new path and decided to discover what the world of industry had to offer. She accepted a position with Procter & Gamble, managing several large-scale production teams, and oversaw all aspects of the production pipeline, including marketing, certifying, packaging, and logistics. A “practical study” in management, as she calls it—to Artigas it represented a major shift in mindset: “Moving from a research mentality to a real-world business meant a move from taking the time to find the *one right answer* to a very specific question, to finding the best possible answer for the demands being made *today*.” During her five years at P&G, she attended business school at night, and received her Masters of Business Administration in Industrial Management, at the University of Ramon Llull in 1995.

At another critical juncture in her career, Artigas decided to take a one-year sabbatical from P&G in 1995. With a desire to have a greater positive impact on society, she went to Costa Rica where she worked with a local Non-Governmental Organization. Her role there was to investigate local employment practices, to help ensure they were compatible with environmental sustainability. Her experience in Costa Rica strengthened her passion for helping local businesses succeed, which she brought to her next position—as the Director of Entrepreneurship & Business Incubators with the Barce-

lona City Council. In a post-Olympics recession, when the Internet was still new, Artigas combined her managerial and engineering skills to create the Barcelona Virtual Incubator—a virtual learning and collaboration platform based on Lotus software, to help facilitate the review 1,000+ new business proposals she received every year. In 1998, Artigas was invited to present the Incubator project at an annual conference organized by IBM.

Artigas discovered something while at the Barcelona City Council that would propel her career forward—at the time, new tech startups had no access to special funding and very limited access to small loans. Promoting an initiative for a venture capital fund for tech startups led to her next move—as CEO of the Ericsson Innova Venture Capital Fund in 1999. During her time at Ericsson, Artigas acted as an advisor for technology companies throughout Spain; she was often the only woman on these advisory boards and acknowledges that there is still a gender gap on company boards in Spain, even today. Reflecting on this experience, Artigas believes that for women in business today, she doesn't feel that there's a glass ceiling being imposed the way it's usually thought of, but rather—it's a question of whether women are ready to pay the cost to break through the “glass.” For Artigas, the question she asked herself was: “Are you willing to pay the price, and when, in order to go as high as possible?”

In 2003, Artigas chose to leave industry, and took two years off to focus on family and enjoy the experience of having her daughter. The time off also led her to reevaluate her career and future. This is when she decided to start her own company in 2006—Synergic Partners, a consultancy firm focused on big data and data science, where Artigas continues as president and partner today. Synergic Partners was recently named “Spain’s Big Data Pioneer” by 451 Research, an industry analyst company focused on enterprise IT innovation, placing it among the top consultancies in the field. What’s been central to the creation of her own company is Artigas’ belief that both women *and men* deserve a balanced family and professional life. She feels that at the same time that women have been partially excluded from the business world, men have been prevented from spending equal time with family. Today, she encourages her male employees to claim for themselves time in their private lives (by taking paternity leave, for example), just as she encourages women to claim their

professional lives, and share equal responsibility for childcare, for example, with their partners.

In looking back on the lessons she's learned over the evolution of her career, Artigas' advice for other women in business comes down to a clear, distilled vision: "Preserve your idea—persist even in the face of challenges." She adds:

"Being an entrepreneur and running your own company is the top ceiling. I see more and more women entrepreneurs that chose this path as their evolution path instead of struggling in companies where they never find the right balance. In my particular case, I see this experience as closing the loop—going back to my roots in maths and science, but being able to capitalize all my experience in the business world."

Francine Bennett

CEO and Co-Founder, Mastodon C

Trustee, DataKind UK

B.A., Maths and Philosophy, University of Oxford



The first in her family to go to college. A mentor with The Young Foundation, an organization whose goal is to inspire and encourage young people to succeed in a complex world. A leader in using data science to challenge complex humanitarian and environmental issues. Francine Bennett is CEO and Co-Founder of Mastodon C, and a Trustee of the first DataKind chapter, in the UK, where the idea of "data-for-good" drives projects in homelessness, childhood poverty, and access to education, to name a few. With a background in math and philosophy, extensive experience in local government, and a clear dedication to improving the lives of others and our environment, Bennett brings a unique skillset to the field of data today.

Growing-up, Bennett was inspired by her math teachers, who fueled her interest and excellence in the subject. Her favorite thing about math, even today? The problem-solving aspect. “I love that there are abstract structures, and you can learn something unexpected by understanding these structures,” says Bennett. From the ages of 16-18, she focused her studies on math, english, art, and music, where she studied the saxophone. She explains that math and music are inherently related, in that both deal with conceptual structures. “Math is about creating and making sense of structures. Music is about playing with structures,” says Bennett. At the University of Oxford, she chose to focus her studies on math and philosophy—two fields she sees connected by one bridge: logic.

“Both fields of study are useful to me now because at its core, programming is about thinking through problems, just the same as with math and philosophy. Especially in data science, you have to consider: what are the implications of these facts that we know? And what’s the impact of what I’m doing,” says Bennett.

After completing her Bachelors degree, she took off for Japan and a completely fresh experience—with no background in teaching, and no experience speaking Japanese, Bennett embraced the adventure and spent a year teaching young people English. When she returned to the UK to begin her PhD in math, she was six months in when she faced a life-changing moment—the realization that she wasn’t happy doing research. With a change of direction that would set her on course for her work in data today, Bennett chose the field of local government, where “there were lots of interesting problems to solve,” and “where I could have a positive impact on people’s lives,” she says. Bennett served for three years as an advisor on how to integrate technology to improve social services.

She then went on to spend three years at Google as a business analyst, followed by 1.5 years at Ask.com, where she was recruited as the Operations Director for Europe. Within these roles, Bennett learned how to write code, use data to inform business decisions, and ask precisely the right questions to help businesses grow and develop. This background in business and skill for asking the right questions led the local charity Off Centre to chose Bennett as a board member. Off Centre serves as a resource for youth in the London borough of Hackney, and provides services such as counseling and art and drama therapy to help resolve issues young people face; it also helps them gain access to government services. Ultimately,

her work with Off Centre is what led her to DataKind—she attended a DataKind hackathon in London and was so impressed with how they were using data to solve problems that she worked to initiate their first chapter in the UK, which launched in April 2013.

In that same spirit of adventure, Bennett launched and self-invested her company, Mastodon C, with co-founder Bruce Durling. A team of agile big data specialists, Mastodon C builds products that use open source technology and the skills to help organizations realize the power of data they already have. Her inspiration to create the company came from realizing that there was lots of good open source technology available to work with data in the cloud, but it wasn't being used for development in important fields such as environmental and government data—it was mostly restricted to Internet companies. Bennett envisions new opportunities for working with data in the *physical* environment, where the Internet of Things is just the beginning of this type of application. One area she's exploring at Mastodon C, is the ability to track energy usage and environmental data, such as temperature and humidity, in buildings, to help improve existing work environments and build more efficient buildings in the future.

With a background centered on service and passion, Bennett offers a single piece of advice that is both timeless and far-reaching, to women and men alike—she says:

“Changing directions—whether it be changing your field, your mind, or your decision—is always OK. Nothing is totally final. If you follow your interests, they’ll lead somewhere good. And if they don’t, you can change your path then!”

Michele Chambers

President/Chief Operating Officer, RapidMiner

MBA, Duke University

BS, Computer Engineering, Nova Southeastern University



The sunny island paradise of Saint Croix, U.S. Virgin Islands, is a place most only dream about as an ideal vacation getaway. For Michele Chambers, this is her hometown—a place where early island experiences imbued her with strong values that have remained for a lifetime; values like transparency, accountability, and self-motivation that have carried over into the corporate ecosystem:

“Island life is a paradigm for the ideal size of a corporate team,” according to Michele, “when you have the opportunity to lock eyes, you have to get along with each other—that’s an environment where results materialize and learning occurs.”

With diverse roles in corporate and tech leadership, and three books on analytics, Chambers knows how to bridge the divide between technology and the marketplace, keep pace with change, and master new skills and tools along the way.

When asked if it is better to be a generalist or a specialist these days, Chambers replied without a flinch—a generalist:

“As a data scientist, you are more of a generalist. You have a combination of skills in math, computer science, and business knowledge. These skills can be applied to hard-core development, business analysis, or entrepreneurial endeavors. This gives women a lot of flexibility as their careers and lives unfold ...For instance, let’s say I leave to have a child; after coming back to work, what I can do is arrange for a reasonable work schedule to balance my current life situation. I can manage a team, or I can move into data visualization; even if I don’t know that much about it at the beginning, I can pick it up easily, armed with what I have already accomplished. I can continue to develop my career based on where I left off, with

something less demanding time-wise; maybe these new areas are less deep tech-wise, but they keep me close to the field that I know and love, with the flexibility I need for where I am now.”

Chambers considers the data scientist’s role to be one that relies on deep insight and communication. She feels that strong communication skills have been essential to her role in helping businesses build unique analytic roadmaps for their enterprise. Chambers points out that today, analytics are very different as compared to 10 years ago; “first generation” analytics were very dependent on descriptive statistics—looking at and describing what had already happened. While many more advanced analytic techniques, such as data mining, machine learning, predictive analytics, artificial intelligence, simulation, and optimization have existed for quite some time, they were rarely adopted in large enterprises, mostly due to lack of skills and computing power. Now, Chambers says “with the advent of new low cost computing technologies, these more advanced techniques are getting deployed to predict outcomes with great precision, at scale which realizes so much more value for the corporation.”

Chambers believes that “now, the key to unleashing value for an enterprise is finding the right talent,” and that women who enter the data space benefit not only from hard skills, like engineering, but additional skills or degrees, such as an MBA.

Finally, Chambers offers some tips for women working in the data space: speak up and ask for what you need, in terms of pay, work environment, and schedule, and be aware of the language you use and accept from others. “Never use self-deprecating language, she notes, “after all, if you don’t believe in yourself, who else is going to believe in you?”

Camille Fournier

Chief Technology Officer, Rent the Runway

MS, Computer Science, University of Wisconsin at Madison

BS, Computer Science, Carnegie Mellon



According to Camille Fournier: “There are only two hard problems in computer science: cache invalidation and what to wear on Friday. We’re solving both.”

Fournier is CTO and former head of engineering at Rent the Runway—a startup fashion company that gives women access to the latest in designer fashion, through a rent and return model. She’s responsible for overseeing the technology architecture, strategic development, and engineering operations for the storefront software, mobile technology and apps, warehouse operations, and reverse logistics platforms. Fournier is also an Apache ZooKeeper Committer and Dropwizard framework member of the Apache Software Foundation Project Management Committee.

What intrigues Fournier about her role at Rent the Runway is the opportunity to solve interesting problems and drive business value with data-driven solutions. According to Fournier: “there are plenty of data professionals out there who can tell an interesting story about the data, and even sometimes a very fun and cool story to support a branding story. But, if it doesn’t add value, what’s the point?”

For Fournier, whose interests come together in software, fashion, and business, her work is all about adding value. She explains: “A huge reason why I’m here at Rent the Runway is that it was an amazing business that was succeeding despite its technology; it was a huge opportunity for me to help something be successful that was going to be successful anyway.” She explains that whether for the customer, the business, or the shareholder, what you’re doing has

got to add value; adding that there is so much data out there that if you don't know where to focus, you can simply drown in the endless ocean of data and get totally lost. Fournier believes that you have to know what is useful and practical, and what is not. An expert in applied data engineering, she relies on her solid foundation in computer science to help her identify what's important in her work.

When asked about the line between data science and data engineering, Fournier notes that there's a bit of an overlap between the two that can get "fuzzy." She explains that strictly speaking—being a data scientist can depend on the company where one works and how that company defines it, what it means to them, and how they configure their internal teams. Sometimes, there are cross functional teams that work together to collaboratively and collectively produce the results associated with the term "data science", whether that be programmatically or by analyzing spreadsheets. In addition to a baseline knowledge of math, science, and programming skills, Fournier feels that what is crucial for anyone working in data is the domain and product knowledge (i.e., what is intrinsically going to add value to the business).

As a field for women, Fournier comments that admittedly, it's tough: "I know too many women who have dropped out of the industry due to working conditions—they are bad for everyone, but women are like the 'canaries in the coal mine.' The startup world is really broken," she continues, "women don't get as much venture capital funding, VCs gravitate to male-dominated startups, so we end up reproducing the same male-dominated structure that existed 50 years ago. However, I believe that our industry is still somewhat immature, so these things may get better over time."

To help resolve the underrepresentation of women in data and tech fields, Fournier feels it would be useful to expose kids to coding in elementary or middle school, so that everyone has equal access to it, and the opportunity to see if it's something they're interested in pursuing. Fournier also acknowledges the importance of female leaders in tech, and the role they can serve as mentors for other women in the field. Fournier herself serves as a mentor for other women in tech, and spoke of the importance female mentors played in forging her career path; in fact, she fondly remembers her grandmother pointing her toward the tech field, telling her how much promise it held for women in the future.

Additionally, while studying at Carnegie Mellon, Fournier noticed an environment that really supported women in science and engineering. Entry requirements were not so rigid around coding that they dissuaded those without that experience, and support programs and on-campus peer groups were all very helpful for making connections with others and getting help when you needed it.

Carla Gentry

Founder and CEO, Analytical Solution

BS, Math and Economics, University of Tennessee



In an impressive *tour de force* of reverse engineering, Carla Gentry labeled herself the “data nerd”—a designation she wears proudly as a badge of courage. Having been teased in school as “nerdy,” Gentry did what came naturally to her—she turned a negative into a positive. Her timely personal rebranding coincided with the advent of big data and the corresponding need to make sense of it all; Gentry loves to ask the question: “Data: what can it do for you today?”

Carla explains: “I’m a data nerd who loves to help companies glean insights from their data. I am able to take huge complicated databases, decipher business needs and come back with intelligence that quantifies spending, profits, and trends.” Carla sees herself as a liaison between an IT department and senior executives, with a keen ability to explain in common sense terms what the customer needs and wants, and what the data is saying. Gentry believes in framing her work as a contribution, and making it visible to executives; doing so comes with significant consequences: it elevates the role of the data scientist to that of a liaison, and makes it a role with business authority. In shining a spotlight on data, Gentry raises the profile and importance of data as a valuable asset that can help drive decision making, thereby making the role of the data scientist strategic.

Gentry started her academic career as a “nontraditional” student—at the age of 27, she was a single mother with two small children, as she began her studies at the University of Tennessee at Chattanooga. She followed her natural gravitation toward math, statistics, and economics and persevered by remembering a significant childhood lesson from her parents: “You learn by doing; live your passion; there are no restrictions except those you place on yourself, and although it might be harder to be recognized as a woman, you can do anything … just keep fighting for everything.”

Gentry uses Twitter to promote the value of data science and spread the word about it being a great field for women, citing the many opportunities to get noticed and have influence. “The sky’s the limit in data science,” she says. “The reward is you can move up the corporate ladder, or open your own business. Have a basic background in math and economics, and it gives you a wide range of options in terms of your profession.”

In her blog *Data Science Is Real: What Can Your Data Do for You?*, Gentry proclaims that “Data scientists are the super stars of the 21st Century.” In her post “Being a ‘Data Scientist’ Is As Much About IT As It Is Analysis” (January 28, 2013), Gentry talks about the unique skillset of the data scientist:

“My definition of the data scientist includes: knowledge of: large databases and clones, slave, master, nodes, schemas, agile, scrum, data cleansing, ETL, SQL and other programming languages, presentation skills, business intelligence and business optimization—plus the ability to glean actionable insight from data. I could go on and on about what the data scientist needs to be familiar with, but the analysis part has to be mastered knowledge, and not just general knowledge ...”

When considering the gender imbalance in the data and tech fields in general, Gentry cited her belief that what would help get more women into the field is to start early, and introduce coding and programming to students at a young age, and to also place a stronger focus on basic skills such as reading, writing, and math. Gentry also believes in the power of positive female role models and mentors, and feels that encouraging children at an early age is especially important in building the confidence and self-worth it takes for them to follow any interest.

Kelly Hoey

Speaker, Strategist, Startup Board Member

JD, University of British Columbia

BA, Political Science, Economics, University of Victoria



Kelly Hoey is a speaker, strategist, and early-stage investor. Her advisory work with startups has taken her into operational roles. In 2014, Kelly was interim Chief Marketing Officer for Cuurio, a startup based in New York City that guides brands through digital innovation.

Hoey is one of Fast Company's *25 Smartest Women on Twitter* and Forbes' *40 Women to Watch over 40* and *Women Changing the World: VC/Entrepreneurs*. And on the New York City-based Alley Watch, Hoey's included on the *20 Awesome People in the New York Tech Scene You Need To Know About*. In addition, she cofounded and was Managing Director of Women Innovate Mobile (WIM), an Accelerator whose mission it was to see women entrepreneurs advance in the tech mobile space.

One of Hoey's top priorities is helping other women, by investing in women-run ventures, and by serving as a mentor to girls who have an interest in the tech field. She was a member of the inaugural class of the Pipeline Fellowship—an angel investing boot camp for women that works to increase diversity in the U.S. angel investing community, and create capital for female entrepreneurs.

Kelly's investment philosophy is broader than rewards and financial return; to her, it's about time, relationships, collaboration, and how to allocate limited resources. She explains: "Investment for me is the how, where, when, and why I choose to allocate limited resources, such as my money and time." Her motto puts it succinctly: "invest in the change you want to see." The change Kelly is interested in seeing and helping bring about is gender diversity:

"I am focused on investing in women, and investing in women early. If we want to see women founding the types of companies which IPO on the New York Stock Exchange or are included in index funds, we need to invest in women-founded ventures as friends/family/seed investors. We need to invest our money, time, contacts, experience in those companies before they have made it. Encouragement is nice, but a check is a lot nicer."

According to the Diana Report, *Women Entrepreneurs 2014: Bridging the Gender Gap in Venture Capital*, in 2011–2013, “15% of the companies receiving venture capital investment had a woman on the executive team; in contrast, a prior Diana Project study in 1999 revealed that businesses with women on the executive team received less than 5% of all venture capital investments.” The study also found that “companies with a woman CEO only received 3% of the total venture capital dollars, or \$1.5 billion out of the total of \$50.8 billion invested during 2011–2013.” This disparity in investment allocation stands in stark contrast to the disposable income power of American women. Per Hoey and the “*She-economy. A Guy’s Guide to Marketing to Women*”:

- 45 percent of U.S. millionaires are women
- 48 percent of estates worth more than \$5 million are controlled by women
- Women age 50+ control net worth of \$19 trillion, and own more than three-fourths of the nation’s financial wealth
- Women will receive 70 percent of inherited wealth over the next two generations
- Women control 85 percent of all consumer purchases.

As part of her effort to bring about greater gender diversity, Hoey is passionate about bringing more women into the tech fields through the use of positive role models. She considers role models and mentors to be especially important throughout all life stages:

“You can’t be what you can’t see and girls with an interest in technology need to see women who have these roles, discover the career paths of those women and understand what career possibilities are open to them... I started the ‘#womenwhotech Google+ Hangout On Air series in 2013’ to do just that. Part of the shift in our mindset with women in technology is seeing those women as experts with a point of view on technology (not simply as having a point of view on women in technology). This is likely why I am so thrilled to be the “CTA” (Chief Technology Ambassador) for Geek Girls and STEM initiatives, for the YWCA of New York City and am

committed to getting as many role models in technology, in front of the girls participating in their Girls Geek Club.”

In terms of the role that early education can play in getting more girls interested in technology, Hoey suggests a cross-disciplinary method that approaches “technology education as language, creativity, and problem solving.... Create an understanding of what technology creates and underlies—from websites to architecture to fashion. Make it relatable to the products and services they [girls] use every day or aspire to have.”

At the higher education level, Hoey again stresses the role of “meaningful internships which provide skills, career guidance and introductions to influencers.” In the workforce—whether women are working for companies or starting their own businesses—Hoey believes that what they need most is “a network of insider relationships.”

Cindi Howson, Vice President of Research, Gartner, Inc.

MBA, Management Information Systems, Rice University

BA, English, University of Maryland



Last summer, the National Center for Women & Information Technology (NCWIT) issued a chart showing women’s share of undergraduate degrees in five STEM categories ([Figure 1-1](#)). The most dramatic line in the chart is the percentage of women in computer and information sciences—although it spiked in the early 80s, the trend turns abruptly downward and has declined ever since (aside from a brief upturn in the mid-90s).

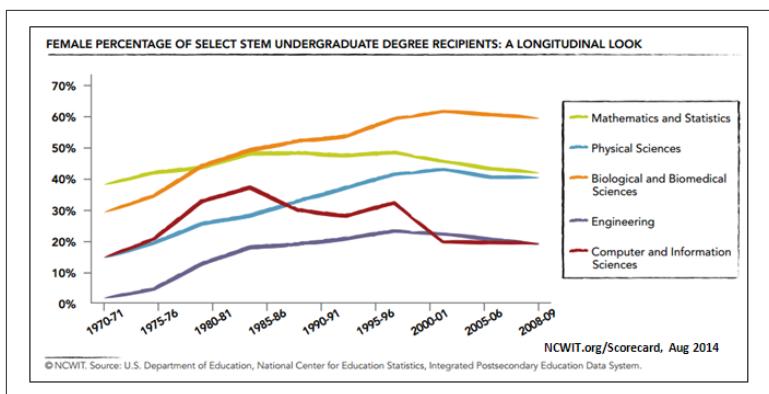


Figure 1-1. Women's share of undergraduate degrees in five STEM categories (1970-2009)

Cindi Howson has completely bucked this trend—she entered the field in the mid-1980s, as things were beginning to take off and has stayed with it ever since. Even more remarkably, Howson didn't plan on a tech-oriented career in the first place—in college, she was on her way to becoming the next great American writer until a detour led her to explore the functionings of local area networks, databases, file recovery, the intricacies of Lotus 123, and FOCUS. In a fateful move, Howson lost some important documents on her computer and her journey to recover them led her down a path toward an interest in technology. Then, when her first employer started a data warehouse project, she was involved in the selection of BI tools, and her journey into BI and data really began.

Years later, Howson returned to school to round out her grasp of business concepts; she completed her MBA in Management Information Systems at Rice University. This additional education prepared her for the launch of her business: the BI Scorecard. According to Howson, “I never had a dream to own my own business—I was simply a new mother at the time trying to juggle work and family and thought being self-employed gave me more control over my time. It's allowed me to do more and have greater success than if I had stayed with one firm. I'm blessed to have a job that I love, that brings me to fascinating clients, and places around the world.”

Fast forward to today, and for 12 years, Howson has been managing her own business consulting with customers on BI strategy, tool selection, and best practices; publishing in-depth product reviews

on BIScorecard.com; writing for *Information Week*, and authoring several books including: *Successful Business Intelligence*, *Unlock the Value of BI & Big Data*, and *SAP Business Objects BI 4.0: The Complete Reference*. She's been a distinguished faculty member at The Data Warehousing Institute (TDWI) and joined Gartner in 2015.

Howson has a big-picture vision of data and analytics—she views them as powerful assets that can offer a competitive advantage to those who use them well. Howson thinks about how data can help companies make the best products, stretch public funding dollars further, and even improve education.

Angie Ma

Chief Operating Officer, Co-Founder, The ASI

PhD, Physics, University College London

Graduate Diploma Law, BPP Law School

MSci, Physics, Imperial College



First, identify what's important to you—whether that be how to make your startup succeed, what career to pursue, how to balance work with family life, or how to improve yourself. Angie Ma is Co-Founder and Chief Operating Officer of The ASI—a London-based startup that runs a carefully structured “finishing school” for science and engineering doctorates. In her role at the ASI, Ma helps guide over 100 PhD graduates a year to find what's important to them personally and professionally, as they transition from higher education to industry. With a background bridging the fields of physics, medicine, law, and data science—Ma's career is driven by a sense of adventure, a commitment to quality, and a desire to provide meaningful service to others.

One of the most influential women in Ma's life also embraced that same sense of adventure—Ma's grandmother was a young girl in

China in the early 20th century, when girls weren't expected or encouraged to go to school *at all*. She and her sisters would take turns going to school—she would go one year, a sister would go the next year, and they would teach each other at the same time. At 18, her grandmother left home with no money and no support, in pursuit of a law degree; Ma remarks, that having achieved it, her grandmother went way beyond her family's and society's expectations, and the value of her grandmother's education is a positive influence that still resonates within her family today.

Rejecting outside influences pushing her to become a pianist or practice law, Ma reveals a history of returning again and again to her passion: science—chosen for its ability to explain how things work, and its capacity to change the world in meaningful ways. Physics, as Ma puts it, is precisely the study of how things work, and she easily relates it to her work in data science, at the ASI today:

“Physics is about understanding the natural world using data. Data science is about understanding the human world with data. The two fields are identical in their goal, which is to develop a deeper mental model of reality. The difference is in the application—what you use the data for.”

After completing her Master of Science degree at the Imperial College, Ma completed her Graduate Diploma of Law at the BPP Law School. While she appreciated the philosophical aspects of studying law, and the perspective it gives her today on any legal issues that come up at the ASI, the most critical thing she learned was that a career in law wasn't for her.

Ma went on to complete her PhD in Physics at the University College London, where she developed miniaturized medical devices for biological applications. Her milestone creation was a 2cm x 2cm microchip with optical circuits, to take the place of a full-size microscope and collect information useful in diagnosing disease. The work required that she build theoretical models and develop a framework for how to measure information received from the microchip. In addition to building models, her experiments also demanded a working knowledge of programming—Ma used both Matlab and Mathematica—both of which are used in industry today, and similar in structure to Python (which she uses now). From her roots in the world of physics, Ma brings a sense of play to her work: “Exploring data is exciting because it alters your perspective, in ways that you may never have imagined without studying the data.”

At the ASI, Ma brings together several key lessons and skills honed during her research experience—one of the most important being the ability to communicate clearly and engage an audience. “Data means nothing unless you have a story that you can tell to explain it,” says Ma. She goes on to explain that a crucial question to ask yourself when working with data is: “How can you relay the story of your data with topics that people care about?” Four years in Toastmasters taught Ma an important answer to this question: always find a way to make the story relatable—and capture the audience’s attention within 30 seconds or less, or risk losing their attention altogether.

Another surprising lesson Ma brings from the world of physics to data science today is the importance for women to see what she calls *the real* problem. She admits that during her PhD research, a lot of any gender-related constraints she faced were a result of her own thinking.

“When you see a problem as a gender problem, you’re seeing it through a tinted glass. You’re not seeing *the real* problem. Instead—if you consider how you can use your own skills to solve the problem, that’s empowering and it brings male colleagues closer because you can ask them for help, and you’re not creating an artificial gender divide,” Ma explains.

When asked for advice for women entering the field of data today, Ma lends a unique and realistic perspective: “Throughout your education, you’re told you can do everything—family and career. Sometimes, it may not be the best advice, actually. In the real world, you have to work with a more realistic framework—being conscious of that will help you make better choices for yourself.”

Neha Narkhede

Cofounder and Head of Engineering, Confluent

MS, Computer Science, Georgia Institute of Technology

BE, Computer Engineering, University of Pune



Growing up in India, Neha's parents made sure she clearly understood that nothing would prevent her from succeeding in any field. This psychological belief in an unlimited future potential, family treatment with complete gender equality and a pro-learning cultural environment in India have all certainly contributed to Neha's success and career choices in later life. These same factors may also shed light on how Neha has pursued her interests in math, computer science and tech with unbridled enthusiasm. A liberating belief system and an independent spirit have shaped Neha's identity with a special combination of entrepreneurship, technology and desire to solve real world problems.

When Neha joined LinkedIn a few years ago, the company, in the throes of explosive growth, had started to encounter multiple instances of scaling problems in its core database, social graph system, search engine and data warehouse. The result was that none of the systems that stored data could keep pace with the variety of data generated by its growing user base which has since expanded to 329 million users worldwide and which adds users at a rate of more than two new members per second.

To accommodate the growth needs, Neha and her team focused much of their time on transitioning to distributed data systems to scale LinkedIn's infrastructure horizontally and incrementally and in bringing in Hadoop to scale the data warehouse. This transition was, per Neha, "powerful and inspiring. It was powerful in the sense that the learnings from applying computational power to the data allowed building very cool new products." In addition, collecting

business and activity data as well as operational data like error logs and performance metrics allowed generation of all this as a set of feeds in Kafka. What started as a way of copying high volume real-time data ended up as a different way to think about the data altogether. Neha became convinced that they were onto something big. And they were. Since its creation by the LinkedIn team, Kafka has seen widespread adoption across thousands of companies like LinkedIn, Uber, Netflix, Verizon, Goldman Sachs, Salesforce and many more.

Ultimately, the team donated the Kafka code to the Apache Software Foundation and they wound up starting a brand new venture, Confluent, to bring Apache Kafka and real-time data to the enterprise. Kafka acts as both a replacement for traditional enterprise message brokers and a way to sync data between different systems and databases. Neha uses an analogy to explain Kafka's functionality: "Kafka acts like a company's central nervous system. If you consider Hadoop like the long term memory or brain that stores huge amounts of data, then Kafka acts as the nervous system that collects billions of sensors in the form of bytes of data for later access." Now with the team at Confluent, Neha "is solving a very key problem that all large companies face, which is collecting data about everything happening inside a company and allowing applications and systems to react to events in real-time."

While by profession a distributed systems engineer, Neha has worked with many data scientists and opines about the ideal skill set combination. "Math and statistics are absolutely essential, coding is an either or situation although knowing how to write code grants authority, speed and freedom of expression," she maintains.

Neha has a keen desire to encourage other women to enter the field and makes several recommendations. Pipeline improvement has to be a focus of attention. For this, providing role models and examples of success via organizations like ShePlusPlus and Technovation Challenge is working successfully with definite results. These groups concentrate on accelerating girls along the way and are really needed for early exposure, modeling positive behavior and proactively dismantling the stereotype that girls are not suited for tech. On the job, she advocates "support from a dedicated mentor and mentoring others is both satisfying and great two-way street learning experience."

Claudia Perlich

Chief Scientist, Dstillery

PhD, Information Systems, New York University

MS, Management Information Systems, New York University

MS, Computer Science, Technische Universität Darmstadt/Technical University Darmstadt

BS, Computer Science, University of Colorado



Detecting breast cancer—early. Classifying the genetics of the yeast genome. Solving a myriad of business and societal problems. From computer science and information systems to entrepreneurship, academia, the business sphere, and parenthood: welcome to the world of algorithms, data, and assorted tales of Claudia Perlich.

Perlich is Chief Scientist at Dstillery, where she tackles the complex feat of analyzing billions of web user interactions, and develops algorithms to score and match cohorts of prospective audiences with major marketers. In addition, Perlich teaches in the Stern MBA program at NYU, has published over 50 scientific articles, holds multiple patents in machine learning, and has won awards and competitions in data mining and related topics. A conscientious and devoted mother, aspiring cellist, and equestrian, Perlich still finds time to give back to the next generation of women.

Perlich spoke candidly about the field of data and its opportunities: “The market is hungry for data scientists. You can even dictate the terms of your engagement. Practices in data science are less formal and rigorous than in the engineering field … you can have flexibility, if you need it.”

Perlich explains what it is about data science that captivates her: “I like the notion of creativity—it’s often around solving the problem

to the algorithm. Intuition is a vital piece on the data side; you often can tell if something is wrong with your data; more often than not there is something wrong with your data, and there is a notion of intuition needed as you review the data—it can't be solved technically..."

Why is this so? In the evolving field of data science, Perlich sees the following set of challenges: "First, the primary challenge is to understand the (business) problem. It's important to have an eye-level ... conversation to make sure the problem is being advanced correctly or that the right question is being asked at all. Next, what data is available to solve the problem from a practical usability standpoint or can be had to solve the problem. Thirdly, what algorithms and what sequence of steps need to be taken to tie it all together." According to Perlich, last but not the least of all of the challenges of a data scientist is convincing business partners of the correct solution; she finds that using clear visualizations are most helpful for explaining solutions in a way that makes sense to everyone in the room.

In terms of career advice, Perlich reflects on her own journey: "My career is a direct result of my natural curiosity and a deliberate delay of early commitment to a specific field." She continues, "I have emphasized flexibility. In my career and professional evolution as a data scientist, I have been agile and have taken advantage of many diverse opportunities and options presented later on in the marketplace."

She strongly recommends that others also avoid early specialization, and instead she believes that early education should have a stronger focus on "structural thinking, the thought process behind algorithmic thinking, and taking the steps needed to solve problems later in life."

Kira Radinsky

Co-founder and CTO at SalesPredict

PhD, Computer Science, Technion-Machon Technologi Le' Israel

BS, Computer Science, Technion-Machon Technologi Le' Israel



Guided by her mother at the age of six, Kira Radinsky wrote her first line of code. Born in the Ukraine, Radinsky immigrated to Israel with her parents as a preschooler. As a child, she delighted in playing with puzzles and riddles and putting things together—a harbinger of future work in conditional probabilities.

In thinking about the inspiration for her work, Radinsky explains: “There are little puzzles you need to solve, and it’s hard to do manually, so I wondered—is there a way of doing this automatically? I see interesting patterns, and I want to know why. I want to be the Indiana Jones of predictive analytics. Today, I feel that many decisions in the business world are based on gut feeling, but we have the information we need to make better decisions based on data—this is my focus.”

Early in her career, using knowledge readily available on the Web—historical records, decades’ worth of news stories from *The New York Times*, social media posts, and many other large-scale sources of data—Radinsky has designed a methodology with Eric Horvitz, codirector at Microsoft Research in Redmond, Washington, that determines when and where disease outbreaks might occur. Their system maps historical data—applying computational power that allows for finding patterns and making certain correlations and inferences about the future. Droughts in Angola and Bangladesh, for example, foretold of cholera outbreaks—thereby linking the weather, disease, and death.

These days, Radinsky is at work in her own company—SalesPredict, where she applies to business the same analytical thought processes

previously used for the weather and healthcare challenges. SalesPredict uses algorithms to help companies find and keep future customers by providing predictive insights into who is most likely to buy, who will buy more, who will churn, and why. Radinsky explains: “If predicting the future can be done, it really would have transformational value for many industries (not only health and healthcare), but for the benefit of mankind and businesses in general.”

According to Radinsky, data science is a multidimensional field where doing well requires many skills. In addition to those usually mentioned—math, statistics, machine learning, coding, etc.—Kira adds business knowledge, “as it allows one to ask the right questions and understand the problem you are trying to solve. Over the long run, coding, and especially being emotionally mature, helps you solve business problems more efficiently.”

From the vantage point of having grown up in Israel, Kira observes important cultural differences between Israel and the United States. “There is greater gender equality in Israel, perhaps stemming from male and female compulsory military service,” she surmises. Also, she notes that in the United States, more women tend to disappear from the work force to have children and don’t return. Also, she does not recall seeing women afraid of math in Israel; in fact, girls were often better at math than boys. “I think it’s more of a cultural influence imposed on girls/women that says they aren’t good at math. Essentially, men and women have the same abilities.”

Majken Sander

Business Analyst & Data Nerd

Business Developer, T. Hansen Gruppen A/S

Business Degree in Financial Analysis, Project Management and IT, Vejle Business School; PBa Math and Art, CVU Jelling



For Majken Sander, working with data is an art as well as a science. In data, she finds a creative outlet that leads to better business decisions. Sander is a Business Developer at T. Hansen Gruppen A/S, a retail supplier of equipment and spare parts for bicycles, scooters, vehicles, computers, and more, in Denmark. She is a self-proclaimed “realist”—interested in what can be measured and proven, and as she puts it: “leaving no path within analytics, visualization of facts, and the language of math unexplored.”

At the young age of 16, Sander began business school, with an entrance into marketing, economics, math, and IT, followed by an advanced focus on business intelligence; she also went on to study project management. But before all of this, her introduction to the field of computer programming began one day at age 10, when an older boy in the neighborhood got a computer—and she couldn’t touch it. The common idea back then was that “programming was for boys” and if something had to do with data or math, you had to be a boy to participate. This only served to fuel her interest, and she went on to both learn how to program and excel in her math courses. It was only later she discovered: “I realized I wasn’t patient enough to be a programmer—I wanted to see the real-life relevance of my work.” Today, she uses her knowledge of programming to “speak the same language” as the programmers she works with, as they work toward common goals.

Sander completed business school and one of her first roles was as a business consultant at Dolberg Data. Her role there involved advising clients on how to run a business, develop business processes, how to make different departments work well together, and how to determine which questions to ask when making business decisions. She vividly recalls that at first, it felt like she was playing ball on the men's field: "I felt that I could be feminine, but not too feminine...I felt that I always had to prove my worth." She later went to work at Hans Tørsleff management systems, and it was at this time that she recalls the first significant milestone in her career: when she felt that a client took her seriously.

"During a project that was off-track, I had to tell an older, male client (and a manager) that he was causing part of the problem by stalling the decision-making process. At the time, I was in my 20s (about half his age), and I was faced with the challenge of telling him his role in the problem. He didn't like it much, but in the end, he thanked me for helping him realize the issue, and he paid the bill."

After working several years as a business consultant, analyst, and project manager in a male-dominated field, Sander took some time off to have children, and at the same time, to re-think her career. Well-practiced in asking tough questions, she asked herself what she really wanted. Her next move was to become a math and science teacher at the Skovvang School, where ironically she discovered why she believes so few women pursue careers in math and science:

"One issue is that a lot of examples used in math are geared toward the general interests of boys (for example, 'How fast does the car drive?'). If we brought in more feminine examples, we could meet girls in their territory, instead of expecting them to get interested in what the boys are interested in, *and* learn the math." She adds: "Math is about numbers, but we use words to talk about it—words the kids have never really thought about, so that creates an added layer of challenge. My focus was to work with kids on their language skills, to bring language into math in an understandable way. Helping the kids learn about words that are used in math."

In the business-world, up until this point, Sander felt "in-between" genders in the office—because of her analytical nature, she didn't feel that she quite fit-in with the other women, and while she got along with her male colleagues, at the end of the day, she wasn't "one of them" either. Being a teacher allowed her a break—to "be myself, without the effort", as she puts it.

After two years teaching, she reentered the business world, rejuvenated and ready to use her talent for analyzing data and business processes at T. Hansen Gruppen A/S. Sander's particular passion is for providing what she calls "data-driven decision support," so that business leaders use data—facts—to make decisions, rather than gut instincts, which she notes is all too common. For example, in her work analyzing data on the sales of spare car parts, Sander often sees that "...a team might estimate that just because they sold 1,200 items of a certain car part two years ago, and 1,000 a year ago, they'll project 800 will be sold in the current year. But, if you look at the data for the cars on the road, for example, if fewer people are driving the car that uses that part, you can make a better prediction, using data to know what to expect, and make decisions based on more realistic expectations."

Today, Sander finds it's easier than ever to be herself in the business world—she acknowledges that some of this comes with age, and some from perspective. Her advice to other women in business reflects a wisdom clearly gained from experience: "Don't feel you have to give up your femininity to be in business. At the same time, you have to consider that you're entering a male-dominated playing field...acknowledge that your male colleagues see the world differently than you do." Sander adds that women bring a diverse perspective to the workplace, and it's important we be aware of its value. She believes that it's in the combination of both men and women's attributes that we can excel in business, together.

Gwen Shapira

Software Engineer, Cloudera

BS, Computer Science and Statistics, Tel Aviv University



Gwen Shapira got an early start in computer science. As a young student in Israel, Tel Aviv University welcomed her as an auditor in

their advanced programming class. She took all of the exams and embarked on the path to her career in the IT industry.

Equipped with a solid foundation in statistics, programming, and database administration experience (Linux and Oracle), Shapira is today part of the engineering team at Cloudera that is responsible for designing and implementing scalable data processing solutions built on Apache Hadoop. She is a committer of the Apache Sqoop project, active contributor for Apache Kafka, and author of *Hadoop Application Architectures* (O'Reilly). Shapira loves troubleshooting and confesses that when she's working with computers she is "*totally in the flow*."

For Shapira, career management has meant selecting the right jobs for her skill set and expertise, preferring deep technical roles. "Women are often gently directed away from technical roles into positions requiring soft skills like project management. It can be tempting, but I found that I'm happiest when working on technical challenges," says Shapira. Shapira is an Open Source developer, a former consultant and an expert on designing data pipelines in Cloudera's data ingest team. The products her team builds enable customers to move data to the Hadoop platform while ensuring accuracy and integrity of the data. In a typical day, Shapira designs and codes new features for Apache Kafka, writes documentation, discusses use cases and solutions with customers, and prospects and perhaps even works on presentations to evangelize her software projects.

In terms of cultural differences Shapira has noticed between the United States and her native Israel, she notes that in Israel, it's not unusual for women to work and have a family life as well. "It's part of the culture. It's just accepted and taken for granted. The whole country is like that. Women just work—like everybody else."

Laurie Skelly

Data Scientist, Datascope; Course Designer & Instructor, Metis

PhD, Integrative Neuroscience, University of Chicago

BA, Neuroscience, Trinity College



During graduate school, Skelly began researching the neural processes behind empathy and psychopathy, connecting the dots between behavior and emotions using machine learning algorithms. Now, as a lead course designer for the Data Science Bootcamp at Metis, a newly launched organization offering a 12-week training program for data scientists, Skelly knows firsthand which combinations of tools, techniques, concepts, and skills are needed by the professional data scientist today. “At Metis, we train for real life conditions. Our curriculum centers on technology, but since real life includes working with others, reacting quickly, troubleshooting, and communications, we design the projects accordingly,” says Skelly.

Skelly’s students learn from the experience of other professionals, work independently and in teams, adapt quickly to the unexpected, and respect deadlines—all in an effort to mimic the real-world working conditions that await them upon graduation. Presentations and public speaking provide opportunities to practice making convincing arguments for future employment interviews and for winning the support of executives on the job. Learning how to communicate sophisticated ideas in a digestible and memorable way is another equally vital component of each student’s success.

Given her multidisciplinary background ranging from academia and research to business consulting and course design, Skelly favors classifying data science as a “generalists” field. As such, she believes it leaves one open to develop their career from a base of technical skills, and expand into business and real-life challenges.

When asked about her views on early education, Skelly commented that she'd love to see the math and science curriculum de-coupled from goals of bygone days and better tailored to modern use cases and interests. For example, she recommends de-emphasizing geometry and trigonometry in favor of concepts from statistics or linear algebra.

In the work world, Skelly would also like to see more males advocating for and supporting women. In her mind, when male coworkers actively promote their female teammates as equals, they clear away unnecessary debris and help everyone to work at their peak ability.

Kathleen Ting

Technical Account Manager, Cloudera

MS, Computer Engineering, Santa Clara University

BS, Computer Engineering, Santa Clara University



As a Technical Account Manager (TAM), Kathleen Ting helps Cloudera's strategic customers deploy and use the Apache Hadoop platform in production. Acting as the customer's technical advisor, she is a technical expert with a passion for customer management. A key aspect to success in this role is persistence: in managing expectations between the customer and Cloudera's product development teams, in forming a relationship of trust with the customer, in anticipating customer needs, and in acting with agility in unexpected situations.

Thinking back to how she wound up in technology, Ting recalls that in high school she attended a week-long engineering camp held at Santa Clara University (SCU). This experience, and the professors she met there, convinced her to major in computer engineering as well as attend SCU. Furthermore, it was at SCU that she met two of the mentors who have helped shape her career.

At SCU, Ting had the opportunity to select Marie Wieck, an industry stalwart at IBM, as a mentor through MentorNet. Encouraging her to continue on for an advanced degree, Wieck afterwards passed Ting's resume into IBM, leading to an interview, and Ting's life-long dream to work at IBM was fulfilled.

After Ting worked at IBM on the mainframe for a few years, it was her former SCU Dean of Engineering, Dan Pitt, who piqued her interest in big data. Although she lacked open source experience at the time, Ting persisted in trying to interview at Cloudera. Eventually it was her unconventional offer to work for free that landed her an interview at Cloudera, where she started in early 2011 as their first Hadoop support engineer.

At Cloudera, Ting sought out Apache Software Foundation (ASF) member Arvind Prabhakar, who became her mentor related to all things Apache. In fact, her first conversation with Prabhakar was around the need for an Apache Sqoop mainframe connector by which to easily move data from the mainframe to Hadoop. Ting credits this mentorship with leading her to become an Apache Sqoop Committer and Project Management Committee (PMC) member. Drawing from her work at Cloudera, Ting is a frequent speaker at data-related conferences as well as a published coauthor of the *Apache Sqoop Cookbook* (O'Reilly).

Appreciative of how mentorship has propelled her career, Ting tries to do her part in shrinking the tech gender gap by volunteering with the Society of Women Engineers as well as by speaking to young women at the She's Geeky conference and to incoming SCU engineering freshmen.

In addition to using mentoring to shrink the tech gender gap, we need to foster creativity in schools, to build the mindset of learning from failures (rather than being discouraged or ashamed), and to set mandatory programming requirements to build familiarity with and exposure to technology at an early age.

Renetta Garrison Tull

**Associate Vice Provost, Director of PROMISE, University of Maryland
Baltimore County (UMBC)**

PhD, Speech Science, Northwestern University

MS, Electrical Engineering, Northwestern University

BS, Electrical Engineering, Howard University



Dr. Renetta Tull is a recognized expert in women and minorities in education, and in the STEM gender gap—both within and outside the academic environment. Dr. Tull is also an electrical engineer by training and is passionate about bringing more women into the field.

From her vantage point at the University of Maryland Baltimore County (UMBC) as Associate Vice Provost for Graduate Student Development and Postdoctoral Affairs, Dr. Tull concentrates on opportunities for graduate and postdoctoral professional development. As Director of PROMISE: Maryland's Alliance for Graduate Education and the Professoriate (AGEP) program for the University System of Maryland (USM), Dr. Tull also has a unique perspective on the STEM subjects that students cover prior to attending the university, within academia, and as preparation for the workforce beyond graduation.

Dr. Tull has been writing code since the seventh grade. Fascinated by the Internet, she “learned HTML before there were WYSIWYGs!”, and remains heavily involved with the online world. “I’ve been politely chided in meetings for pulling out my phones (yes plural), sending texts, and updating our organization’s professional Twitter and Facebook status, while taking care of emails from multiple accounts. I manage several blogs, each for different audiences ... friends, colleagues, and students.”

Dr. Tull has been involved with many strategic initiatives, including the following:

- Liaison for Institutional Collaboration at Universidad Metropolitana in San Juan, Puerto Rico
- Co-PI for the ADVANCE Hispanic Women in STEM Networking Conference
- Co-PI Collaborative Research on diversity programs with the Quality of Life Technology Center (Carnegie Mellon University/ University of Pittsburgh)
- Leader, “Women in STEM” project for the Latin and Caribbean Consortium of Engineering Institutions (LACCEI).
- Invited speaker for programs at MIT, Cornell, University of Maine System, Society for Hispanic Professional Engineers (SHPE), National Society of Black Engineers (NSBE), American Indian Science and Engineering Society (AISES), The National GEM Consortium and others
- Former Board Member of the Northeastern Association of Graduate Schools
- Former Vice President of Operations for an emerging technologies firm

Dr. Tull believes that women thrive with support and a network. She recommends organizations such as [Latinas in Computing](#), the [ACM Richard Tapia Celebration of Diversity in Computing Conference](#), and the National Science Foundation’s [Louis Stokes Alliance for Minority Participation](#), a grant-based program that provides student scholarships.

Dr. Tull works to increase the psychological sense of community and professional development opportunities for graduate students in Maryland. Through her involvement with PROMISE, a local program in Maryland that is one of eight National Science Foundation AGEP programs in the United States for underrepresented STEM graduate students and postdocs. Programs include the Summer Success Institute (SSI), Professors-in-Training (PROF-it), Dissertation House, Research Symposium, and others. According to Dr. Tull, these programs have contributed to an increase in applications, enrollments, and graduation rates of underrepresented graduate students in STEM fields in Maryland.

Dr. Tull sums up her thoughts about some of the changes and initiatives that need to be implemented across the STEM fields:

“Women can be more involved with computing fields, and certainly, opportunities for ubiquitous computing limit restrictions for engagement. However, women need to know that they are invited to the table; they need to see images that show that they are not excluded and they need regular and continuous access to computing environments early, within the K-12 years. Many local, regional, and national groups are doing that. These groups need exposure within the schools, so that girls will take advantage of them. Further, the national media can contribute to a solution by showcasing images of girls working with computers, and women in computing. Some leaders are talking about making computer science a core standard as early as elementary school, in the same way that math and reading are taught at early ages. Building proficiency, and acceptance at a young age, can go a long way. This acceptance must then continue throughout college, graduate school, and the career.”

Hanna Wallach

Researcher, Microsoft; Assistant Professor, University of Massachusetts Amherst

PhD, Machine Learning, University of Cambridge

MS, Cognitive Science and Machine Learning, University of Edinburgh

BS, Computer Science, University of Cambridge



In speaking with Hanna Wallach, she proudly cites a point of great distinction: “I’m probably the only person on the planet who has appeared in both Linux Format and Glamour magazines.”

Another example of bridging different worlds—Wallach works in research at Microsoft and also has academic responsibilities at the University of Massachusetts Amherst. At UMass, Wallach is a core faculty member in the recently formed Computational Social Science Institute, where she develops new machine learning methods for uncovering fresh insights about the ways in which people inter-

act. She collaborates with political scientists, sociologists, and journalists to learn how organizations—often those underlying the U.S. political system—work in practice by analyzing publicly available data. Her research contributes to machine learning, Bayesian statistics, and in collaboration with social scientists, to the field of computational social science. Her work on infinite belief networks won the best paper award at AISTATS 2010, and she's organized several workshops on Bayesian latent variable modeling and computational social science.

A Glamour Magazine “35 Women Under 35 Who Are Changing the Tech Industry” honoree, Wallach says “at its core, my work is really all about using fancy math and fast computers to learn about social processes, such as those that underlie the U.S. government.” She adds that what enables her cutting-edge work is the massive quantity and diversity of data now publicly available. “We’ve had computers for a while. We’ve also had some of the math. But now we have massive amounts of data. What were once hypothetical theories about people and society are now being validated (or disproved!) by large-scale, data-driven analyses. This is extremely exciting.”

Is a researcher and computer science professor also a data professional? “Yes, absolutely,” says Wallach, “because data science brings together people with a wide variety of skill sets and backgrounds.” Moreover, she believes that in order to succeed in the field, you need to be open-minded, interested, and continually learning—you need to have a “*growth mindset*.”

Wallach enjoys a strong network of professional women and men, who are at the ready to provide support, suggestions, and advice as role models, friends, and colleagues. Committed to giving back and supporting the cause of helping other women, Wallach has cofounded three groups that focus on helping the next generation, including the annual Women in Machine Learning Workshop (which is currently in its ninth year!).

From her vantage point, Hanna sees a few fixes to the educational system that would help bring more women into the data and STEM fields. She believes it’s critical to involve girls in science as early as possible (waiting until college is way too late, in her opinion). In addition, she encourages visible role models, promoting diversity and inclusion, and networking. Wallach strongly believes it’s important to embrace and seek out diversity: “There’s even considerable

research about how teams with diverse viewpoints end up creating better ideas and products.”

Alice Zheng

Director of Data Science, Dato

Post Doc, Carnegie Mellon, Auton Lab and Parallel Data Lab

PhD, Electrical Engineering, emphasis in Communication, Computation and Statistics, University of California, Berkeley

BA, Mathematics and Computer Science, University of California, Berkeley



Alice Zheng is Director of Data Science at Dato—the Seattle-based startup that was awarded at the Startup Showcase at the Strata +Hadoop World conference in Fall 2014. At Dato, Zheng demonstrates her expertise in machine learning algorithms and applications—she loves playing with data and building tools that enable others to play with data.

How did this come about? Alice recalls being in elementary school in her native China, where math studies begin early, and bringing home a math grade that was below expectations. Her parents, both aeronautical engineers, were convinced she could do better. And after some intense tutoring from her father, she did, thereafter to remain at the top of all her math classes.

To Zheng, data science is essentially about understanding data: “data science is about making cohesive sense of those many many tiny pieces of information. It’s about detecting patterns that can then be used to help people make better decisions.” She adds that while many academic disciplines focus on the method, “data science focuses on the object of analysis. It’s a science (and art) all about the nails, not the hammer—and I like that!”

Zheng explains her feeling that the core of data science is problem solving:

“I love solving problems, and I love solving problems through data. But even more fundamentally, I love gaining fresh perspectives of the world. There are as many perspectives as there are people, and data is as amorphous, as flexible, as mysterious, and as potentially impactful as human beings. ... I delight in the process of carefully deriving meaning and actionable information where it was previously dominated by chaos and confusion. It’s almost a secondary benefit to then be able to make decisions or produce outputs (like making recommendations of what else someone can buy), based on the understanding I gain. The insights are what I prize.”

Zheng believes that the key to success is removing obstacles, even one's own inner obstacles: “it's our perceived limitations that limit us, not our innate abilities. To encourage girls to take up math and science, I think we should be working on getting rid of all of their perceived limitations about themselves. This is subtly different than building confidence. Confidence is the positive booster, whereas perceived limitations means negative self-talk. As we look for positive feedback to build confidence, we should also be eliminating the sources of negativity.”

Zheng also has a refreshing perspective on failure—she affirms that failing at something at the beginning is not a sign that one shouldn't do it. While failure is certainly uncomfortable, Zheng believes it's not where we *start* that should be the determining factor. She explains: “As human beings, we are very good at gauging position (where we are in terms of expertise level), but we are not so good at noticing velocity (how fast we are moving up the chain). It's the latter that will determine the level of expertise we will eventually reach. This is why it's important to ignore those uncomfortable failures at the beginning, because it may take us a while to ramp up and maybe even longer for us to notice the change. ...”

Margit Zwemer

Founder and CEO, LiquidLandscape

MFE, Financial Engineering, University of California, Berkeley

BS, Mathematics, Stanford University



Growing up in a family where both parents and grandparents had advanced degrees, it came naturally to Margit Zwemer to have a strong motivation toward higher education, and even to be comfortable with math and enjoy a challenge. Zwemer's grandparents emigrated from Hungary; her grandfather arrived with a PhD in chemistry and rebuilt his business here in the United States. The value of a scientific education was instilled in Zwemer at an early age. In fact, it was her father who taught Zwemer how to code when she was still in elementary school.

Zwemer's early professional experience was as an options trader (a "quant") at Goldman Sachs, and later with Société Générale. When she joined tech startup Kaggle, she rebranded as a "data scientist." From having worked with machine learning algorithms and the open data community at Kaggle, Margit later developed these interests and techniques further in the financial data visualization space.

In early 2013, Margit cofounded her own financial technology startup, LiquidLandscape, a company focused on real-time human/data interaction using high-volume financial time series data. "Founding a startup is decidedly an exciting time with ups and downs, predictably, but having the right cofounder is a huge part of building a strong company," Zwemer shares. Having her own business allows Zwemer to use all of her skills and interests—to apply math and engineering to solving real world problems, and to use charts and visuals to enable understanding by her business constituents.

Zwemer sees data visualizations “not just as pretty pictures to summarize the underlying data but as a tool for interacting with and manipulating a live data set and searching for patterns and anomalies that may not be easy to identify in the raw numbers.” At Liquid-Landscape, Zwemer is now “exploring a broad range of visualization environments each of which has pros and cons for understanding certain aspects of the data. We use both traditional two-dimensional visualizations like line charts and heat maps to show the interactions between a small number of variables and three-dimensional screen graphics that allow the overlay of additional streams of data so that users can visualize the overall state of the market.”

The company is currently exploring immersive data environments in virtual reality. Zwemer adds that what’s important for us now is “being able to integrate these visual environments into a single core data pipeline; it’s a huge leap over traditional ways of monitoring and visualizing data and consequently making decisions, optimizing performance and reducing risk.”

Zwemer sees data science as an emerging field and a hybrid one, that paradoxically, lowers the risk associated with starting one’s own business. “From my perspective, being a hybrid field, it lessens the economic risk of looking for work because it is multifaceted. As one evolves and develops throughout one’s career, it’s always good to have a multiplicity of tools in the tool kit. Data science provides that. And since data scientists are in demand, we can take more risks, career-wise, and that is very acceptable, even valued as a badge of courage. Data science enhances your risk tolerance with others and with your personal life and yourself.”

About the Authors

Cornelia Lévy-Bencheton is a communications strategy consultant and writer whose data-driven marketing and decision support work helps companies optimize their performance.

As Principal of CLB Strategic Consulting, LLC., her focus is on the impact of disruptive technologies and their associated cultural challenges that open up new opportunities and necessitate refreshed strategies. She concentrates on big data, IT, Women in STEM, social media and collaborative networking.

Ms. Lévy-Bencheton has held senior marketing and strategy positions in well-known financial services firms, is currently on the Board of The Data Warehouse Institute, Tri-State Chapter, (TDWI) and the Board of the Financial Women's Association (FWA). She earned her MA from Stanford University, MBA from Pace University and holds several advanced certificates from New York University.

Shannon Cutt is an O'Reilly editor focused on content related to data. She has more than seven years of experience developing and managing books, video, and other media. When not reading about data, she's hiking all over the East Bay Area and caring for any stray animal who comes her way.
