

## **Ben Nelson**

Data Scientist at Chan Zuckerberg Initiative

*“Meta, a research discovery tool for the biomedical sciences”*

As science advances, it becomes increasingly challenging to keep track of the scientific literature. Over one million papers are published each year in the biomedical sciences. Scientists can't build on each others' work unless they can discover it, and better tools for absorbing the scientific literature could enable every scientist to do better work. The Chan Zuckerberg Initiative's Science Tech team is building Meta, a research discovery tool that provides a faster way to understand and explore biomedical science as it evolves. We use machine learning to present users with personalized feeds of relevant papers and preprints -- both at the level of broad research fields and at the level of an individual's specific research interests. I will discuss the current state of Meta and my role as a data scientist to understand what value researchers are getting out of the platform.

## **Dan Caselden**

Distinguished Security Engineer at Gigamon Applied Threat Research

*“Brown Dwarfs or Bad Guys? Leveraging Machine Learning to Find Oddities”*

Mining astronomical catalogs with machine learning has yielded thousands of new brown dwarf candidates. We took the very same methodologies for finding brown dwarfs and applied them to finding and ousting malicious actors in real world networks. This talk will outline both the astronomy and security initiatives and highlight findings.

## **Genevieve Graves**

CEO at Eye0, Inc.

*“Start Up, Consultant, or Big Company?”*

Our academic preparation doesn't give us many opportunities to explore different types of real-world employment. Nor does it do a particularly good job of helping us introspect about what we actually want or need or care about in a job, other than the opportunity to do intellectually-engaging work---we all want that! Since becoming a Data Scientist, I've had the chance to work at several start ups at various stages, pre-Seed to "Growth", as both an employee and as a consultant, and to compare notes with various friends and colleagues who work at large companies. I've done hands-on data science as an individual contributor, I've managed data science teams, and I've been a core decision-making executive. I now run an independent consulting firm, working with clients across a range of business verticals and with varying data science and data engineering needs. In this talk, I will discuss and compare the context of all these different types of jobs---everything that *\*isn't\** the actual data science---and how to self-interrogate about which context is the best context for YOU.

## **Jessica Kirkpatrick**

Director of Data Science at Kobold Metals

*“Treasure Hunting with Machine Learning”*

KoBold Metals is a mineral exploration company that uses machine learning and other scientific computing techniques to increase the ethical supply of critical materials (i.e., Ni, Co, Cu, Li, Pt, and Pd) needed to build electric vehicles and personal electronics. Many new mines will be needed if electric vehicles are to become widespread, but before new ore deposits can be mined,

they must be found. I will talk about how Kobold uses geological, physical, chemical, and satellite data sets to search for new mineral deposits.

### **Julia Silge**

Data Scientist at RStudio

*“Data visualization for real world machine learning”*

Visual representations of data inform how machine learning practitioners think, understand, and decide. Before charts are ever used for outward communication about a ML system, they are used by the system designers and operators themselves as a tool to make better modeling choices. Practitioners use visualization, from very familiar statistical graphics to creative and less standard plots, at the points of most important human decisions when other ways to validate those decisions can be difficult. Visualization approaches are used to understand both the data that serves as input for machine learning and the models that practitioners create. In this talk, learn about the process of building a ML model in the real world, how and when practitioners use visualization to make more effective choices, and considerations for ML visualization tooling.

### **Taka Tanaka**

Director of Data Science at Radish Fiction

*“Building and Managing a Data Team in the Year 2020”*

In July, I joined Radish, a serialized fiction startup with 50 employees, to build and lead the company’s data science and analytics practice. In this talk, I’ll share what I learned from hiring and managing a team at a new company, completely remotely. In particular, I’ll talk about connecting with applicants during the recruiting process; onboarding new teammates; building a supportive and empowering environment; managing emotional health and team morale; and promoting work-life balance—amidst the triple backdrop of a pandemic, a relentless news cycle, and the often hectic setting of a series-A startup.