What I Learned As Pandora’s First Data Scientist

Three years ago, Gordon Rios (https://www.linkedin.com/in/gordonrios) became Pandora (http://www.pandora.com)’s first official data scientist. Since then, he’s seen the team grow to over a dozen strong and become hugely influential in every decision the company makes. Considering how much of Pandora’s service is data-driven — from maintaining its famous Music Genome Project (http://www.pandora.com/about/mgp) to creating even
more ways for people to discover music they’ll love — it’s one of the best examples around of a data science team growing fast and lean to make a difference.

As a keen observer of how Pandora has hired and structured its team of scientists, Rios is a rare source on what works and what doesn’t. And, with the profile of data science on the rise at huge companies and startups alike, figuring out how to maximize productivity in this area is increasingly vital.

In this exclusive interview, Rios talks about the evolution of data science at Pandora, and the **three key lessons** others can take away about how to distribute resources, ensure excellent management, and keep communication nimble to do great science and make it count.

---

**The Fully Integrated Scientist**

If you ask Gordon Rios what he’s passionate about, the two things that come to mind first are science and music. He’s completely fascinated with how people determine what to listen to, why, and how their tastes and habits change. So it makes sense that he’s one of the core contributors to Pandora’s playlist creation, where he works with engineers and other scientists to test and refine playlist algorithms using machine learning.

The key thing to note is that he’s a full-time member of what Pandora calls its Playlist Team. This comes before his identity as a member of the data science team because he is — for all intents and purposes — solely devoted to making playlists the best they can be. He’s embedded with a team of not only engineers, but product managers, designers and others to make this possible.
This isn’t how data science works everywhere. Because it’s somewhat new as a field, it comes in many forms. Some companies have all of their data scientists sitting together regardless of what they’re working on so they can closely communicate. Some even have them working completely separately from the rest of the company, tackling problems and handing solutions over the wall to engineers who take it from there. Others follow a consultant-like model where scientists parachute into projects temporarily to run analyses or answer one-off questions.

At Pandora, they’ve found that embedding scientists with feature-centric teams simply works best — and based on his experience, Rios agrees.

“When a scientist is assigned to a group, we really get to know and become part of that group, and that’s what you want,” he says. “If you’re assigned to the playlist team, you still communicate with other scientists at the company, but you report to the management of the playlist team. You’re part of the program full time.”

Known for making big things happen with relatively scrappy headcount (../..//article/This-Product-Prioritization-System-Nabbed-Pandora-More-Than-70-Million-Active-Monthly-Users-with-Just-40-Engineers), Pandora has made it a core value for people to be committed to doing more with less in every area.

“Our mission on the playlist team is to make sure that music finds its way to a good, receptive audience. We want artists to get listeners, and we want listeners to have the best experience possible. Both of these things depend on getting people to try new music.” And that depends fully on experimenting, collecting data, and designing algorithms that push people outside of their musical comfort zones at the right pace.
So the data challenge Rios and the team face is staggering: Looking only at behavioral numbers, they must determine if someone is happy with what they're hearing, whether they skip songs because they're unfamiliar with them or simply don’t like them, and whether they’re getting bored with their current selection. How they answer these questions impacts individual listeners on a granular level.

“With this kind of project, you need people on operations, engineering, product, and scientists all coming at the problem from different sides, yet with a common vision for the service,” says Rios. “A consultant model would never work. When I first got started with data mining earlier in my career, I often worked as a consultant, and it’s very difficult to make progress on large-scale problems that way. You have to be part of the team to understand all the moving parts.” This is why he recommends keeping scientists on a team permanently unless a program is cancelled or there’s another compelling reason to shift things around. You want them to be fully immersed.

“If you separate out your science team, they can’t do their best work. They'll get bored or spin their wheels.”

The best-case scenario, Rios says, is staffing data scientists that have good engineering skills. When scientists can ship, you save on headcount and you have people with the skills to turn data into meaningful products. This is something to start thinking about from the very start — when you first decide that data science will be core to your startup’s success. Later you can specialize, and that’s important too.
Ideally at the beginning, a company will start with one scientist, much like Rios, who is a Swiss Army knife — someone who can both test hypotheses and code, create an algorithm and deploy it. With experience at Yahoo (http://www.yahoo.com), and as CTO of local search startup Zvents, Rios brought full-stack programming abilities, big data experience, and machine learning expertise to the table. He also had other critical skills that your first scientist needs: the ability to work autonomously, self motivate and be accountable. **First hires in this area are rarely closely managed by anyone, and that’s okay if you have the right person in the role.**

If and when you start building out your science team, you should maintain this flexible attitude, Rios says. Instead of setting your sights only on PhDs and research scientists, you need jacks-of-all-trades who tend to be more interested in practical applications than theory. If you hire people who end up in the weeds, focusing only on experiments without seeing the product implications on the horizon, you’ll build a fraction of what your team could and should produce. **At the same time, hiring PhD’s is crucial:** They have the training to work deeply and autonomously on hard problems that you need to solve to be competitive.

The variable that makes the biggest difference here is proper management, Rios says. “In most cases, good management is about lining up people’s skills to the company’s needs, but with data science, so much depends on having people be both skilled and interested.”

> “When you staff a project with people who are skilled and fascinated by the problem, you get gold.”

The Art of Data Science Management
“Of course there are times when you have to be a trooper and tackle a project that is less than interesting yet critical to the company, but if you have incredibly rich talent, matching them carefully with the right projects has got to be what scientific management is all about,” Rios says. “Being able to consistently do that separates good managers from okay managers.”

Today, Pandora has a director of research for Playlists and Growth and Retention named Oscar Celma (http://ocelma.net/) who embodies everything a company should be looking for in a data science leader. He’s a real hacker who’s also a well-published PhD and leader in the field of music recommendation technology. Pandora plans to build on this managerial model for other areas including Ad Science.

To take a team of scientists to the next level of productivity, managers should ask the following questions:

What is going to move the company in the direction we want it to go the fastest?

Who wants to develop in what way or grow what skill set?

How can we make sure, given answers to the first two questions, that people are put on feature-oriented teams that present healthy challenges?

If you can nail No. 3, “That’s when you get really great results,” Rios says. “When people are put on problems that are too hard for their skill set, they are likely to stall out unless they have a really great mentor to help them out. It’s
possible to stall a whole project that way. But if things are too easy, they can get bored and lose interest.”

“Focus on getting talented people to work together; when you do, you'll see truly inspired results.”

So much of good data science management is about mentor relationships, he says. Even though he recommends for scientists to be embedded within cross-functional teams, partnerships between scientists — allowing them to share best practices, ideas, and solutions with other people they enjoy working with — are what keep very talented people engaged and growing.

As its squad of data scientists has expanded, with a vision that is truly collaborative, Pandora has established several regular communication channels. These include messaging tools like Slack (http://www.slack.com) with channels that go to all scientists and anyone else who is interested. It’s used to ask questions, provide advice, and even circulate academic papers that may be relevant to problems people are working on internally.

Even though they sit separately the majority of the time, the team holds regular meetings and often rallies for lunch to talk about what they’re working on and to have more informal conversations about ideas. A lot of solutions emerge out of these discussions. Slightly more formally, they schedule time to present their projects and findings to their colleagues so they can ask or answer questions, and/or share practices that might be helpful with other experiments. The key has been to let the infrastructure evolve at the same rate as the team, never becoming too onerous, yet maintaining connections so that mentorship is encouraged and work doesn't get duplicated.
According to Rios, a lot of this communication should be about who on the team specializes in which skill sets, even if most people could be considered generalists. “Some people here are classical data scientists, some are better statisticians, and some are better software developers,” he says. Knowing who is good at what helps bring in the right people and solve problems faster.

“Success for a data science program is when people are happy, feeling challenged and fulfilled, and delivering important results. That’s when they are at the highest performance state, delivering the most value,” says Rios. “There are a lot of reasons to bring on junior or inexperienced scientists — they adapt and learn fast — but you better have really solid management and mentorship in place.”

The question is hardly ever if someone is smart enough to handle a particular job, he says. The hiring process for data scientists everywhere he has observed is so rigorous that smarts aren’t the issue. This is doubly true for data science leaders. The real question is whether or not they’re the right fit for the problems being solved, the culture, and to help the scientists who are already there thrive.

“You can always judge a data science manager by how productive everyone is who is working with them,” Rios says. For this reason, the interview process for managers working with data scientists should be extremely challenging and two-pronged:

1. **They should have an informed respect** for everything the other scientists had to know and do to get hired. This includes the ability to create complex models and identify even the most technical features of a given product. They must be as mentally and exceptionally agile as the scientists they will manage, and they must be able to learn new things just as fast.
2. “Everybody has to like them when they come in to interview,” Rios says. “They have to be a culture fit first and then a skill fit. They have to really love the product and know about Pandora and the data challenges we’re interested in solving. It’s tough to get hired into these positions, and it should be.”

“\textbf{The hallmark of being a good manager or collaborator is that everyone wants you involved in their project.}”

Communicate for Optimum Efficiency

“To be an effective data scientist, you have to realize your job isn’t just about the research. You have to quantify and qualify what you do in a way that makes sense across the whole company,” Rios says. “If people don’t communicate, getting results that actually matter to the whole can be frustratingly elusive.”

\textbf{Luckily, when it comes to good communication, your data scientists have a super power.} They can use their ability to quantify things to measure their own success and to concretely convey what they’re doing to the rest of the company. Clear and constant communication is especially important in this area because data science moves so rapidly. You run experiment after experiment, get results, use them to modify the software, and repeat. At Pandora, Rios and his colleagues are constantly hypothesizing what will engage more listeners with more artists. “We suggest changes, prioritize the experiments, and spin them out to a small number of listeners fast to understand if its actually an improvement or not.”
Something that seems as instant and seamless as a song recommendation has a lot of moving parts owned by different people. “We actually have an ensemble of recommendation strategies,” Rios explains. “For example, we could start playing more 80’s music, such as Cindy Lauper or Cher. Depending on a user’s behavior, we can tell if they want more 80’s pop or more recent music such as Lady Gaga or Rihanna. “Behind the scenes, many different things are at work there.”

As you can imagine, this gets complicated when multiple teams are running multiple experiments related to the same experience at the same time. Building and maintaining a sophisticated A/B testing platform is a must. Being aware of what other teams are doing is vital, but it’s really the platform that needs to support multiple working groups.

“There are thousands of things we want to try, but we have to work within the vision for what we want the service to be.”

One of the things that has made Pandora’s approach to data so effective is an appreciation for special skills and how they should be applied, Rios says. It’s no secret which people are really good at recommendation projects, or machine learning or growth hacking user acquisition. Make it standard to foster this transparency every time someone new comes aboard.

“This is why we’re so focused on recruiting scientists who are scientifically curious but also entrepreneurial,” Rios says. “To succeed on this team is to be extremely invested in the area you’re focused on, and equally willing and able to help out where your talents are needed.”
How to Engage Your Users to Build the Product They Actually Want (/article/How-To-Build-The-Product-Your-Users-Actually-Want)