

# Training Unicorns

I saw an article recently<sup>1</sup> that bemoaned the rarity of good “data scientists”. The article explains that to be effective they should combine four distinct talents: firstly a broad and deep understanding of the business; secondly a good feel for how to apply statistics; thirdly an ability to persuade computers to jump through the hoops required to collect and consolidate the source data; and finally the ability to weave together a good narrative and communicate an interesting story. The article claims that finding someone to recruit with this combination of skills is as rare as coming across a unicorn. Now, to me, this list of attributes seems like a rather good summary of what is required. It may also be that finding all of these in a single individual is not particularly common. In my experience they are not as rare as the article tries to imply, I know quite a few people who would match that description, but there again those individuals are memorable exactly because they do match that description.

For the sake of argument let’s assume that individuals with those four capabilities can deliver significant benefits to many different types of businesses. That certainly seems like a very reasonable suggestion to me. In that case any rational businesses that employ those individuals should be very reluctant to see them go, they should do whatever they can to retain them. Thinking about the individuals I know with those skills it is certainly the case that they are unlikely to become available for hire on the open market. So if these individuals are hard to recruit then what should companies do? Obviously they should find ways to grow existing staff, they should encourage people within the organisation that have a subset of these skills to become proficient in the ones they lack. This is exactly why mentoring and “Knowledge Management” are so important for organisations large and small.



But the four skills outlined above are not (in my opinion) equally easy to learn. Deeply understanding the E&P business requires time spent with practitioners, discussions about what their current projects are attempting to achieve and what barriers are making that difficult. It requires devoting time to a wide range of specialists, if you already understand drilling then sit with the geophysicists and geologists, and so on. Proficiency with statistics is about knowing enough of the underlying theory to be

dangerous and applying it, then defending your conclusions, and doing that lots of times. Finding interesting narratives and telling them in a compelling way requires practice (hours spend videoing yourself and reviewing while alone are invaluable in my experience). In contrast the ability to write good software to extract and transform data is a skill I’ve never been able to teach. Perhaps that is because as a software person I can’t remember ever not being good at it. We all know that it is rather hard to catch fully grown wild unicorns, the trick is to identify those that have that promise, and provide the training and practice to help them become the valuable beasts that they have the potential to be.

---

<sup>1</sup> The Guardian article “Data scientists: ‘As rare as unicorns’” by Jeanne Harris & Ray Eitel-Porter is available from: <http://www.theguardian.com/media-network/2015/feb/12/data-scientists-as-rare-as-unicorns>