Deploying R&D analytics at the pace of the data driven business

Adrian Clarke[1]

[1] *Tessella Ltd, Bretby Business Park, Ashby Road, Stanhope Bretby, Burton upon Trent, DE15 0YZ, UK*
E-mail : *Adrian.Clarke@tessella.com*

Abstract
The modern data-driven business is increasingly shaped by the challenges and opportunities that follow from the unceasing growth in the breadth and volume of data. R&D functions have always faced huge demands in the race to be first to market with consumer-personalised products. Addressing these demands leads to new stresses around how to extract the next level of value from data. These stresses are strikingly consistent across different industries and across scientific disciplines, with the same challenges appearing time and again: "How can a fixed central team maintain the skills to cope with the rate of change in scope we face?"; "A proven way to solve this problem must already exist – how do I find it?"; "Progress relies upon taking many iterative steps, so how do I plan my use of the team over the next two months?" Increasingly the challenge is not a lack of mature statistical or analytical knowledge, but how to mobilise it dynamically, in the fast paced world that is modern R&D.

In this paper, I focus upon three key data analytics needs we find in all industries and explain how my organization has evolved to address these needs. In the accompanying conference presentation I will use real world examples to illustrate the points.

Keywords: Big Data, R&D, Analytics, Partnership

1. Introduction

For the last 35 years, my organization has been solving complex problems for clients that span multiple industry sectors. In different forms, data and analytics is at the heart of almost everything we do. Over 60% of our staff hold PhDs in disciplines such as Physics, Chemistry and Mathematics. This academic rigour and ability to quickly understand our client's domains underpins our ability to consistently solve such a broad spectrum of problems. However, in the last decade or so, we have seen a significant shift in the R&D landscape. The volume of data has become unprecedented, though is often difficult to analyse due to growing dimensionality. The distribution of the data within an organization is bewildering and the number of techniques and technologies used to analyse and draw insight from the data has exploded. Our experience gained in responding to these changes in analytics and data science is what I want to share with you here.

2. The Key Analytics Needs

The breadth of our client base is an important aspect of our business. We span many sectors, including Consumer Industries, Life Sciences, Energy, Space and Defence. Interestingly, despite their apparent differences, our experience is that they share the same key issues when it comes to analytics. Fundamentally, our clients know that they are sitting on gold mines of information in the form of existing and future data repositories. However they are finding that realising that potential is hard. The following sub-sections describe the three key analytics needs which, if satisfied, can unlock that potential.

2.1 Speed of Insights

In R&D, speed can often be critical to success. Maybe it's being first to market with a new product. Maybe it's turning around a statistical analysis quickly to allow results to be fed back into the planning of the next field trial or experiment. Answering questions more quickly, by converting data into insight and then knowledge, is absolutely key.

In some cases, clients have told me that they don't bother performing statistical analyses, although they are confident it would save them money or improve their experiments. All because it will take them too long to assemble the data or too long or expensive to get a project approved. Had they persevered, by the time the analysis was completed the data would no longer be current or other changes would have driven a shift in focus.

2.2 Flexible Access to Expert Resources

Answering questions with data quickly is one thing, but in R&D it is hard to plan activities to provide the correct flow rate of questions to match the experts you have available. Add to this having the right data available at the right time and the problem becomes acute. Whilst you could try and view this as a traditional resourcing problem, it doesn't take into account the habit of a research project to take an unexpected turn. Be it exciting discoveries or miserable failures, the research project is by definition uncertain. Having flexible access to the necessary range of people and resources is the second key need. Failure on this front amplifies the speed problem.

2.3 Broad Range of Deep Skills

Fast and flexible access to key resources is difficult enough, but the biggest challenge can be having access to a superset of people with the right skills. R&D in any one domain is a fast changing environment. When you couple that with the challenges posed by the age of Big Data, with new technologies, techniques and possibilities appearing every week, how can R&D teams keep up?

This problem is exacerbated by the fact that sometimes an individual research problem might benefit from introducing a new specialist technique. One week it may be machine learning. The next it could be Bayesian statistics. Often it is a combination of many skills, each one on a short term basis. Having the right depth of skill in a particular capability available in the right place at the right time is increasingly a struggle. This leads to repeated missed opportunities.

3.   Evolving Working Practices

As we recognised the growing impact of the universal needs described in section 2, my organization evolved the way it engages with clients. A feature of our academic based culture has always been close working relationships. Our focus is on understanding business problems and establishing trust based partnerships. This is far preferable to adopting a transactional approach where the client sends their data for analysis and results are returned. This often leads to misunderstandings and little opportunity for innovation with the problem owner. However, traditional engagement models where fixed individuals or project teams are supplied for a set period of time to work against a fixed scope still do not address the core problem.

The Tessella Analytics Partnership was born out of the ideas of true partnership working. A managed service model, it embeds a dedicated analytics partner into the client team, with a focus upon understanding the business problems and a detailed view of their technical challenges. The analytics partner leads a virtual analytics team which is assembled from the Tessella global pool of technical and business domain specialists. The virtual team is managed dynamically, scaled up and down as needed around the pre-agreed average team size, in as little as a few hours if necessary. The skills composition of the team is also dynamically refined, bringing in particular expertise for specific tasks, for example if we want to run a Machine Learning exploratory workshop.

I am an analytics partner with Tessella, working primarily to support global statistics initiatives in the food and drink industry. The embedded nature of my role has enabled me to provide a much more closely personalised service to my clients, backed up by the wealth of expertise across our global organization.

In the conference presentation that will accompany this paper, I will talk through examples of work we have done that have addressed the key issues of speed, flexible access to resources and a broad range of deep skills.

4.   Conclusion

Regardless of the mechanism through which R&D organizations solve the issues I have outlined here, be it using a model similar to the Tessella Analytics Partnership or other means, significant value is waiting to be unlocked from the data that is available to them. The first step in using data to answer questions in a sustainably timely manner is to ensure it is structured, collated and exposed effectively, and that expert resources are on hand to mine for those answers. Structuring their resources to allow flexible access that can match the changing demand profile with which the questions arise is the second step. Finally, facilitating access to an evolving team of experts that encompasses, now and in the future, a broad enough range of analytics skills is the third key to meeting the ever changing challenges presented by the world of "Big Data" within R&D.

Writing these three steps down is easy to do. Implementing them is a significant challenge, one that is shared by organizations across the world and across industrial sectors.