



Successful commercial usage of advanced analytics

by Paul Hewitt, Director and Practice Lead for AI,
Machine Learning and Big Data

Luxoft is proud to partner with Dataiku, and our analytics team have now successfully delivered many Dataiku-based solutions both internally and for clients. In particular, we use advanced analytics not only in an R&D capacity, but also make use of its full functionality in a business context.

So, what are the key success factors for commercially successful implementation of advanced analytics solutions such as Dataiku?

As data scientists, analytics and machine learning is what gets us out of bed in the morning — it's the fun bit! However, successful delivery in a commercial context requires more than the latest technical knowledge.

The three main pillars of commercial success are:

- Context and strategic alignment
- Architecture and integration
- Processes and operations

The place to start is with proof of concepts and proof of values. And there is quite a common pattern in this — all too often, we take a slice of data and produce some models based on that slice of data. However, that initial model is often only a small part of the overall (commercial) picture; getting the model commercially usable in production can be a much larger undertaking.

It's been said that any successful machine learning system has (at most) about 5% machine learning in it. The other 95% is made up of other parts — data and data wrangling in particular.

The data science in the system is fundamental, but to get it into production is more complex than is at first apparent.

Firstly, there's the data journey — which is the overall progression towards AI machine learning at an enterprise level. End users want to get to a point where the data and data analytics can be used to make better-informed decisions, as a consequence of machine learning.

To get there, there are a couple of prerequisites: Data management, obviously, is essential. But the first point that we always encounter commercially is consistent and coherent data models.

Now, this is not necessarily the most glamorous part of the operation, but if you can't get a data commercial model that makes sense to the business, there isn't much point in progressing further.

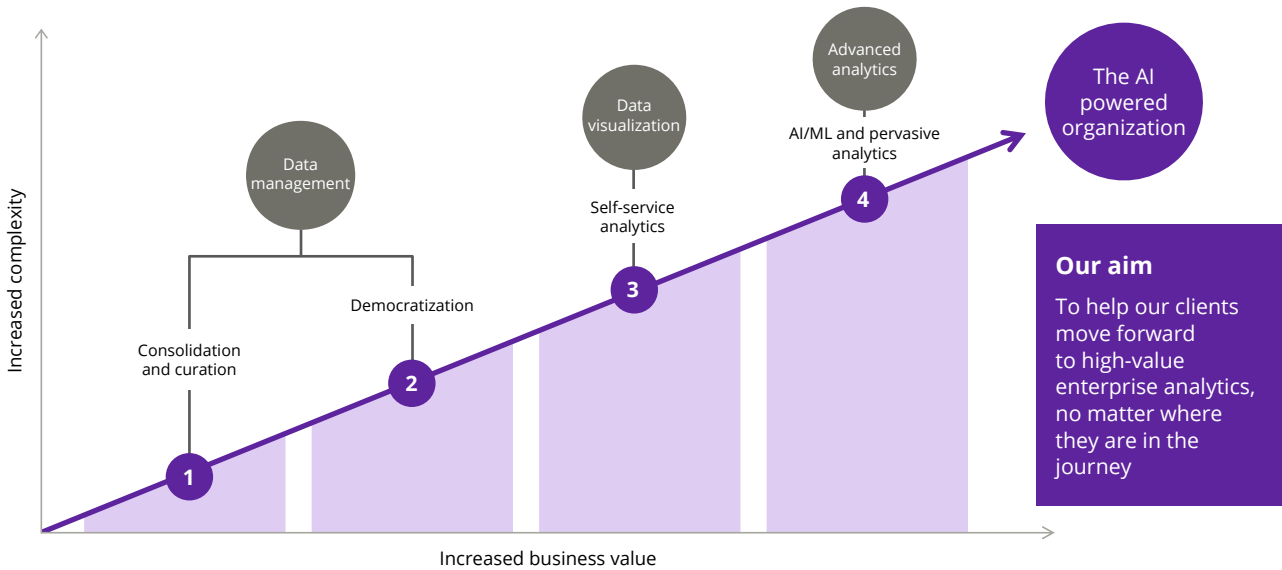
If you make data available to the business, it has to do what it says on the tin. And if you make it available via a self-service type of package, then your end users need to be able to navigate the data model, both easily and successfully.

The data model therefore needs to be complete, coherent and descriptive. The acid test of these attributes is that — if the data model is demonstrated to end users via a self-service package (Dataiku, PowerBI, Tableau or any other similar system) — the business users can navigate and use that data quickly and successfully. If this is possible, you have something with the potential to go viral relatively quickly.

In Luxoft, we've worked on projects where we've been told to expect 15 end users but ended up with 400+ simply because other users in the business heard positive things about it and requested access. This is the power of the data model.

Another way of putting this is — you can't expect to have enterprise-wide analytics if you can't do self-service analytics for the business, which is relatively simple. The data model is, in fact, a gateway to enterprise-wide, commercial-grade analytics for the whole organization. The self-service component, although not directly a technical component, is a fundamental commercial success factor to getting advanced analytics rolled out across the whole organization.

Journey to enterprise analytics







There are four basic pillars of this journey:

- Consolidation and curation
- Democratization
- Data visualization and self-service analytics
- Pervasive advanced analytics

The first few points are foundational, and are often taken for granted, but if you don't get these basic foundations right, you'll be storing up problems for further along the journey.

Detail — the four pillars of the journey

	 Consolidation and curation	 Democratization	 Data visualization and self-service analytics	 Pervasive advanced analytics
What?	Get all the data into one place	Make the data available to users, in the way that they want	Provide curated data and data models, via self-service visualization	Implement enterprise analytics, that can safely be used for enhanced decision-making
Why?	<ul style="list-style-type: none"> • Cheaper for users to connect to one data source, rather than 40 • Easier to do consistent QA 	<ul style="list-style-type: none"> • If people can't get the data into their favorite interface, they're less likely to use it 	<ul style="list-style-type: none"> • Much quicker for users; they know the data best • No need to wait for IT • Many eyes on the data means better reliability • If you can't do enterprise-scale self-service, you can't do enterprise-scale advanced analytics... 	<ul style="list-style-type: none"> • Has potential to add great value to the enterprise • Monetizes the enterprise's (often very valuable) proprietary data • Requires all previous steps to be fully implemented...

It's when we get to the fourth pillar, the advanced analytics, where packages like Dataiku really come into their own.

There are many topical issues within the industry which need to be addressed as a prerequisite to commercial success. These issues — from security to privacy to compliance — are large and complex issues in their own right.

One such family of issues is regulation, compliance and ethics. These comprise not only the different pushes that regulators are currently making, but also include the internal drivers to ensure things are done correctly and ethically.

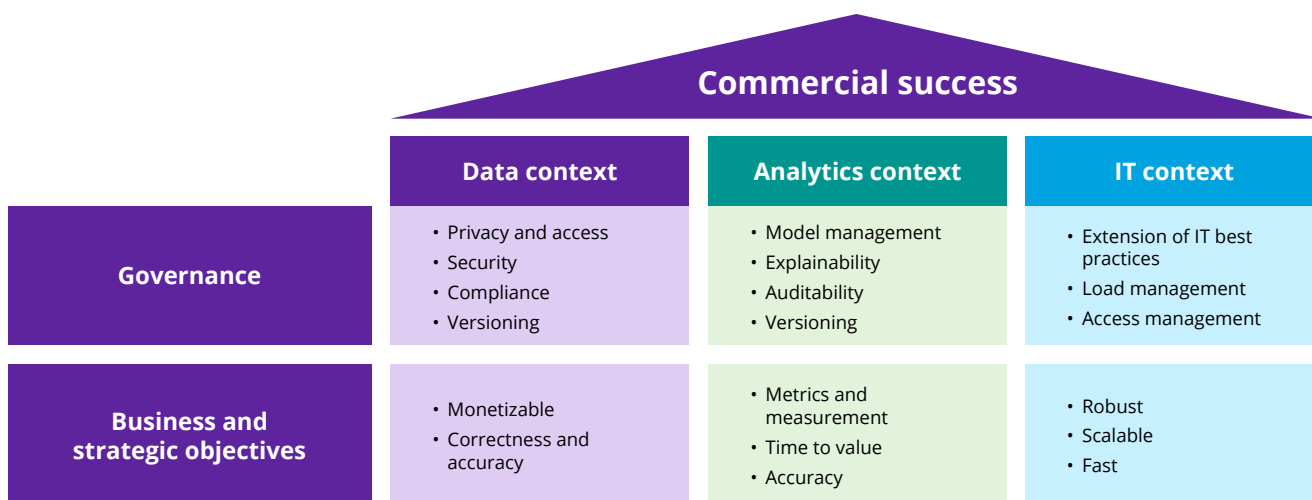
Another topical issue is the direct link between the IT and the data analytics and the ways in which these two fields, in a sense, cross over. We all know that machine

learning operations (MLOps) is becoming a huge topic in its own right.

However, not only is MLOps important, but in fact also, the whole test side is critically important. In recent times, there has been progress in ensuring the metrics, bias and drift are monitored and that errors are managed. But increasingly, analytics is becoming real time.

In real-time analytics, the test side is even more critical — load running and load balancing are essential, and we've seen a lot of progress around issues like setting up large-scale synthetic data sets. With these, you can test the data at scale before actually rolling it out, as well as populating the data with features you want to detect as you go along and testing before implementation.

Some current key issues

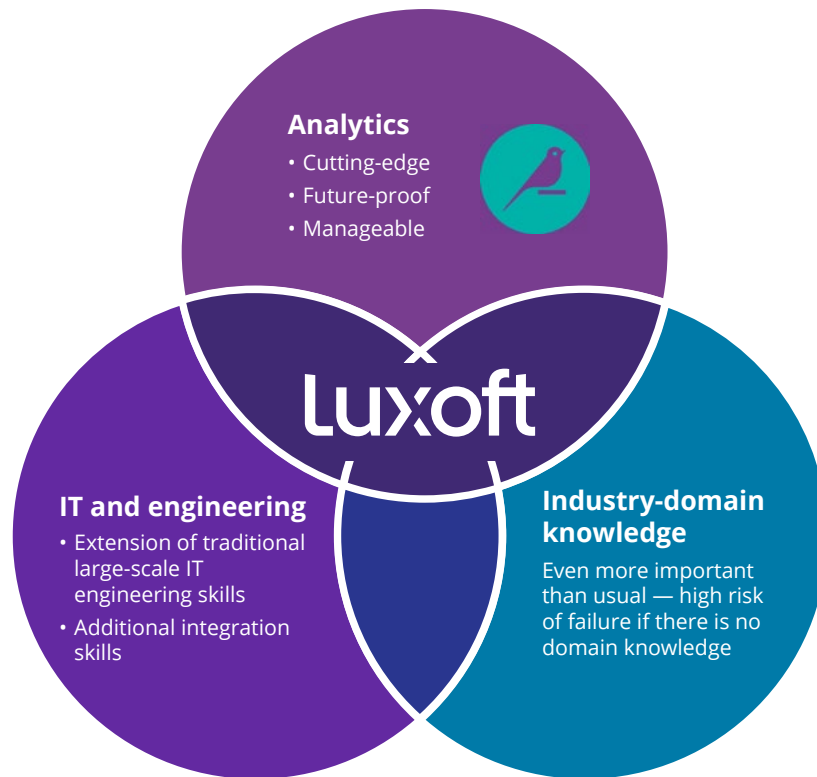


As far as architecture and integration goes, the general architecture is based on widely known models — one data store sitting at the heart of the architecture, with various components layered on top, such as a visual analytics platform. Whether based on Dataiku or others, the principle is the same.

With regards to successful commercial implementation, it's important to focus on the client side — UI and

the client application. The speed of the roll out is very important to end users — they don't want to wait six weeks for a new application. This desire for speed is the driver behind the popularity of low-code solutions which can be quickly and easily developed and integrated. The overall theme is quicker and better integration, with new solutions going into prod quickly for a speedy rollout.

Summary — successful commercial delivery



In conclusion, there's no one key success factor in successful commercial implementation. As well as the basic foundation of technical knowledge in analytics, it's important to combine the engineering and IT side as well as the business processes.

Luxoft regards the combination of analytics, engineering and strong SME business domain

knowledge as prerequisites for successful machine learning and advanced analytics, and for progression along the data journey.

Please **contact us** for any further discussion or insights on any of the topics raised in this paper.

About **the author**



Paul Hewitt

Director and Practice Lead for AI,
Machine Learning and Big Data

Paul leads the Data and Analytics practice for Banking and Capital Markets Consulting EMEA. His role spans all advisory and consultancy in data and analytics, and delivery of all projects from PoCs to large multiyear projects. He has worked with Luxoft for 3 years and has over 20 years' industry experience.

About Luxoft

Luxoft is the design, data and development arm of DXC Technology, providing bespoke, end-to-end technology solutions for mission-critical systems, products and services. We help create data-fueled organizations, solving complex operational, technological and strategic challenges. Our passion is building resilient businesses, while generating new business channels and revenue streams, exceptional user experiences and modernized operations at scale.

luxoft.com