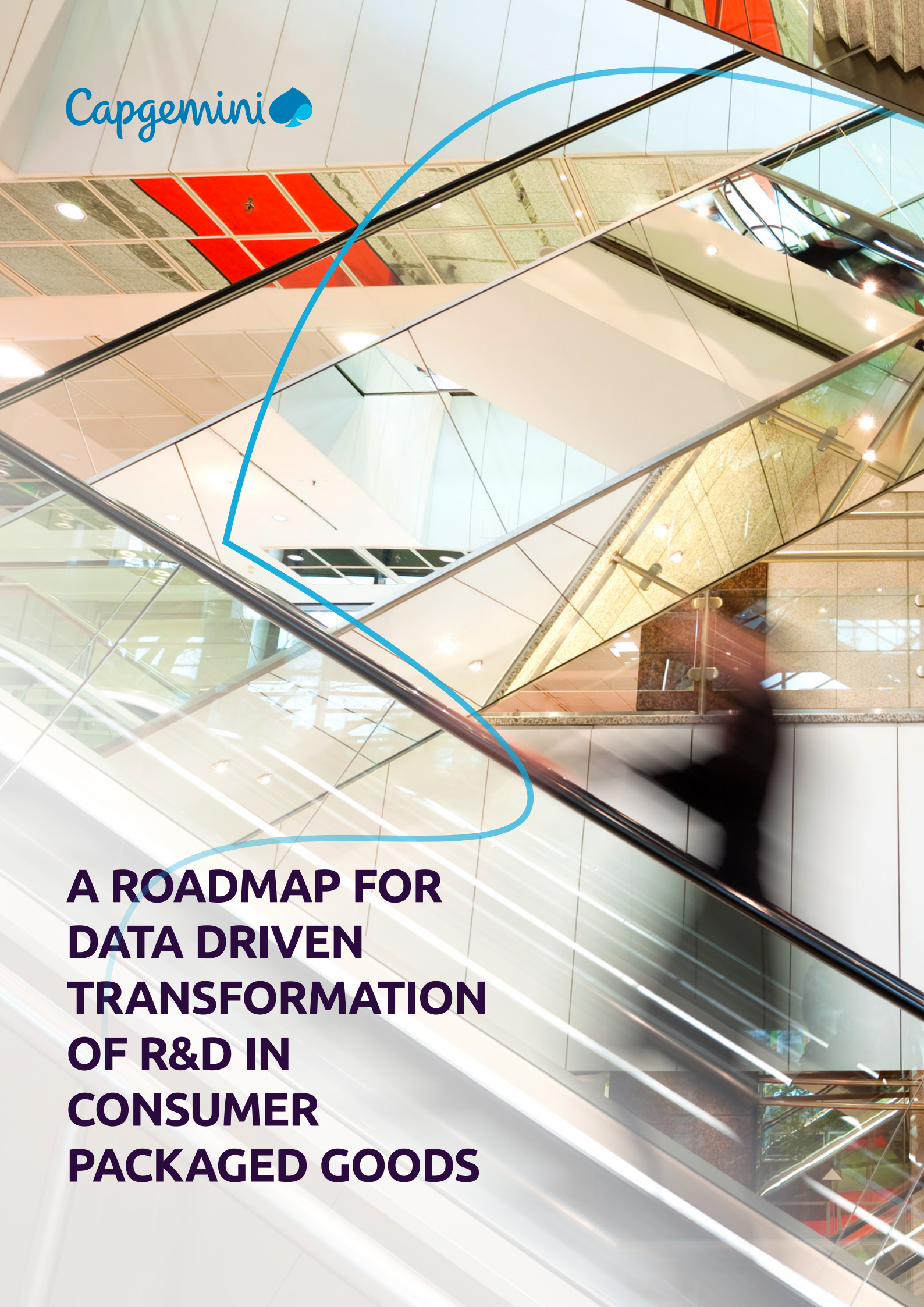


A ROADMAP FOR DATA DRIVEN TRANSFORMATION OF R&D IN CONSUMER PACKAGED GOODS

Contents

Introduction

R&D is the backbone of the consumer goods industries, creating new products, from skin creams, to sauces, to shampoos.

But consumer demands change quickly. Colours and fragrances fall in and out of fashion. Demand is growing for ethical versions of products. Useful chemicals come under scrutiny, whilst others soar in price as geopolitics shift.

Marketing teams want to launch products that resonate with the current mood, and those moods are increasingly personalised across geographies and demographics. Data driven market research teams can spot these changes as they happen.

This puts a lot of pressure on R&D departments to move quickly to bring out new products, or update and personalise existing ones, as well as to change and reuse ingredients. Faster and more targeted R&D would mean shorter times to market. That would allow fast responses to changing supply issues, and allow companies to profit from new products at the start of a trend, not mid-way through it.

R&D takes time, a lot more time than an analysis of consumer trends. But through digitalising many areas of R&D, it can certainly be done much faster than it currently is.

This white paper will explore how.

What do we mean by Digital R&D?

Digital R&D means using data – from experimental data to product images and industry standards – and models to accelerate new product development and update existing products.

These can, for example, predict optimal formulation, how the product performs in situ, or how consumers will react.

The true Digital R&D department has data and modelling at the heart of everything it does, and even links its models into other areas of the business, to optimise the entire product creation and marketing process.

But delivering Digital R&D is not just about building models, it is about creating an organisation built on doing more with data. It is about increasing data maturity by setting up the technology and processes to allow people to rapidly access, collaborate on, and reuse data as needed and establish the skills and culture to be able to quickly harness that to deliver value. It must also be pragmatic, far too many transformation programs become so complicated that the goal becomes impossible and never deliver value.

This is easier said than done in a large and complex business. This is not a single transformation project; Digital R&D can emerge from a whole range of improvements. Every company is different, and just doing some of them maybe sufficient to give a company the boost it needs.

The challenge of Digital R&D is in transforming the organisation so that it is able to capture knowledge, easily transfer it between teams, and harness data to build robust and useful models. This means changes to people, culture, processes, and technology.

How data models drive R&D

- Reducing wasted experimentation time with better in-silico modelling
- Recommending candidate formulations to deliver specific product properties (e.g. taste, smell, shelf-life)
- Predicting product safety to reject formulations early
- Automating regulatory checks
- Providing customer insight & predicting customer behaviours

What's stopping the Digital R&D department

Delivering Digital R&D will require many organisations to make changes to technology, data landscapes, people, and processes, so that innovators, researchers, and data scientists can quickly access the information and data they need, and build models aligned to business needs. Doing so needs to be done in a way that works for the business without causing unneeded disruption to current ways of working.

Some of the key issues that need to be overcome, include:

Data exists in silos

Much of the data that could be useful for modelling – sensory, analytical chemistry, consumer insights – was captured in less digital times, leaving it siloed across different parts of the business. A particular challenge are Mergers & Acquisitions, where data from the acquired company is often never integrated or really understood.

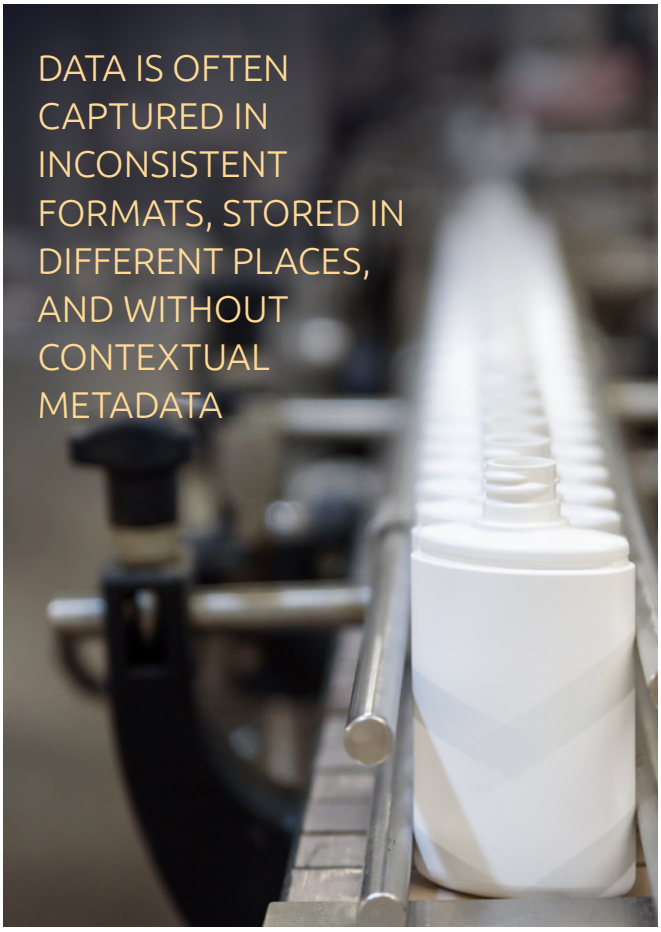
Likewise, considerable knowledge has been captured in personal notes and laptops, or exists only in the heads of scientists. These people will one day retire or leave the company. Knowledge Management programmes designed to capture this information and make it accessible to others can be invaluable in deciding which research avenues to pursue and how to approach them.

Lack of consistent data standards

Data is often captured in inconsistent formats, stored in different places, and without contextual metadata. This makes it very hard to access, and requires considerable work to get it into useable format. In some cases, the cost of this remediation can outweigh the value of doing the project.

Complex IT landscape

Many IT systems have grown organically and lack consistency, leading to many instances of data, some of which have been changed, being stored in different places, including places that are inaccessible. The lack of a 'single source of truth' makes gathering reliable data extremely hard and limits the value of any models or insight. R&D IT often has many complex specialised applications with small user bases, a situation which creates additional complexity for change initiatives.



Lack of support structures for data scientists

Data science teams are often left to get on with delivering a data science project, without support to access the people capturing the data or the people who will use their models, leading to models being built that do not deliver value to the business or which are not taken up.

Narrow skillsets

Data science projects often lack access to the necessary combination of skills and support to see them through to effective deployment. Too many promising use cases never see the light of day because the team can't deploy the most appropriate models or don't have the skills to integrate them into IT systems.

Examples of use cases that make up a data science portfolio

Automating product trials

A toothpaste R&D team found they were manually comparing images of teeth each time they launched a new product, a time-consuming process. We built a data pipeline to ingest image data and automated image analysis using machine learning. This reduced analysis time from weeks to minutes and allowed comparisons across multiple trials.

Reducing human testing

A company making condiments was running costly taste test panels, requiring lots of small batches to be made, most of which would never see the light of day. We developed models that predict subjective opinion (eg 'velvety and smooth'), based on objective product properties (pH, viscosity, acidity etc). This helped reduce tests needed and ultimately shortened time to market.

Predicting shelf life

A drinks company was reducing preservative to suit evolving consumer preferences, which affected shelf life. Even in a lab, shelf life tests are time consuming. Utilising 20 years of shelf life data, we developed machine learning algorithms to predict shelf life for a given formulation, reducing the number of formulations being tested that would then go on to fail and allowing product developers to explore more options without the cost of testing. These have since been rolled out globally for product development teams.

EVEN IN A LAB, SHELF LIFE TESTS ARE TIME CONSUMING

How to progress towards Digital R&D?

Digital R&D is a journey not a project. Different R&D departments have different goals, operate in different environments, and start with different maturity levels.

But there are approaches that consistently deliver value, which we outline below. Organizations need to assess where they are on this journey and how these approaches can be utilised to advance their Digital R&D at a pace that is ambitious but manageable for them.

The best way to advance is by doing. Build a portfolio of data science use cases suited to your digital maturity and use these to learn as you go, whilst delivering real value.

The four stages of the digitalisation journey are:

1. Understand feasibility & value

2. Understand the landscape

3. Build a roadmap

4. Implement & embed

For simplicity, we present the four stages separately, since each has its own goals and requirements. But in reality, there is considerable overlap and all need to inform each other to deliver progression towards Digital R&D.

These four stages inform the choice and approach to the identified use cases. They also ensure that learnings are taken back into the organisation to improve skills, technology choices, and data management.

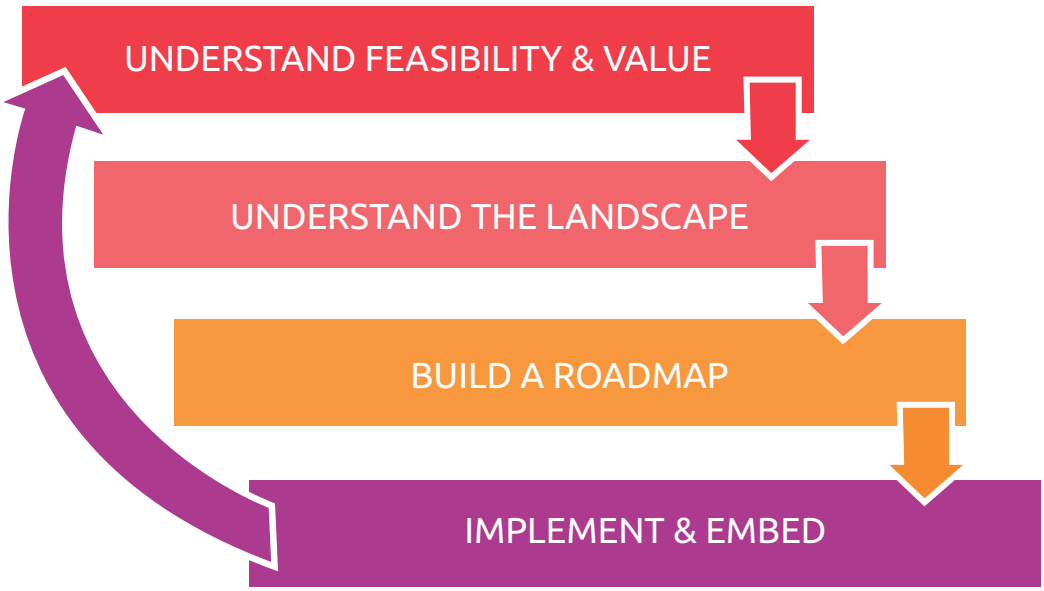
As digital maturity advances, more ambitious projects can be started in a continuous cycle of improvement.

1. Understand feasibility & value

When selecting data science or AI use cases to progress, start with a Proof of Value exercise. This looks at each possible use case and assesses whether, if it were built, it would actually deliver any value.

This is different from a proof of concept which explores whether it can be built. Lots of things can be built, but some cost more to build than the value they deliver, and others aren't worth building because there isn't sufficient quality or availability of data.

Proof of Value helps make the right decisions early and focuses resources in the right direction. It may also identify new datasets that could be captured to deliver huge improvement in predictive ability or identify projects which could deliver value in future once digital



PROOF OF VALUE IN ACTION

Tessella partnered with a major decorative paints manufacturer to investigate whether their data could deliver new insights with real business value. Our 'Art of the Possible' workshop identified which business cases had the most value, which were more likely to succeed, and which could draw on a wide range of the available data sources.

One of the ideas chosen for investigation combined internal and market data sources to get insight into product positioning. Tessella demonstrated that the data was good enough to show the expected macroscopic trends. The investigation also identified key data quality and reliability issues that should be resolved before productionising the models into the client's decision-making processes.

Before launching into a full proof of concept, we proved the value of the existing data by using it to deliver working models that could generate business insights, along with detailed, concrete next steps.

maturity is more advanced and data quality is better, but may not be right now.

One of the most common outcomes of Proof of Value is to identify unexpected holes in existing datasets (or missing datasets) that prevent value being realised. Such problems are one of the most common reasons that data projects become derailed. Spotting this early means the company can understand how to get data ready, before entering proof of concept territory, and eventually full-on projects with expectations and deadlines.

To run a Proof of Value, start with "Art of the possible" workshops which explore potential use cases and ask what they are trying to achieve and what data is available. Data scientists, business and subject matter experts all work together to explore possibilities arising from available data against the intended business goals, and visualise insight. This allows a quick assessment of whether proposed value can be delivered with the existing data. This will identify which offer the most value and should be progressed.

Each project should be supported by what we call a 'Path to Value' to take it from conception to desired output. This output could be anything from answering a question with data, through to building a minimal viable product. Path to Value identifies which systems and processes you need to put in place to collect, clean and pipe data to where it is needed. It may be valuable to rerun Proof of Value exercises at critical points along this path to assess whether new learnings have created new opportunities or spotted roadblocks that need to be overcome.

2. Understanding the landscape

Generate a clear understanding of what your internal structures look like, via a landscaping exercise. This is not about listing technologies, but understanding your level of maturity, by which we mean systems, data, data-science, processes, and strategies. This will help you define what you can do immediately, and where investment is needed to meet longer term goals.

Map what data, systems, analytics, and support are available to you. Capture what data management processes and governance frameworks are in place. Consider existing initiatives that will impact upon digitalisation programmes.

Look at what skills you need to deliver use cases – in data, subject matter, range of modelling options, and software engineering.

Map where new data or skills are needed to deliver promising use cases and how to access them, but be ready to drop anything which will cost more than the value it delivers. Screen data and data sources for the intended application – is it clean, complete, and free from bias, or is more work needed?

In doing so, identify common weaknesses, and develop processes and governance frameworks to ensure future data is captured in a way that makes it easy to find and use, and skills are accessible through hiring, training, or consultants.

Your landscape is a way of helping you understand where data is, and how much effort it would be to use it. This allows informed decisions and investments to be made aligned to use cases or organisational need.

The landscape should not seek to digitise everything and make all data accessible to everyone, since much of that work would have no value. Some data will simply be too hard to access easily or too low value to be worth it. Data without metadata may be easier to recapture than painstakingly sort through. It makes more sense to have an overview that allows you to be agile and invest in ways that will deliver tangible value, than to try to order everything in a complex and evolving system.

Prioritise and target your landscape analysis to get the right level of fidelity in the right parts of the landscape. This takes skill and experience to do.

3. Build a roadmap

Build a roadmap for how you will get from your current state to your desired state, utilising your portfolio of prioritised use cases.

This should establish a series of workstreams — or paths to value - to deliver progress. Each use case should have its own workstream, whilst overarching workstreams should guide use cases, and feed learnings into the organisation.

One overarching workstream should focus on data management, to ensure projects have the quality and availability of data they need. This should inform a wider programme to standardise future data capture and storage – including rules for what format, and what meta data should be recorded, to make data identification easier for modellers in future.

Another overarching workstream should focus on governance, SOP's and working practices.

This will coordinate all use cases and ensure they are delivering in a structured but agile way against key milestones – progressing through data screening, feasibility, development, and real-world deployment. Our RAPIDE data science governance framework provides a useful guide. Once effective governance approaches have been tried and proven across multiple projects, they should be established as an organisation-wide standard for all data projects.

Setting milestones is key to tracking progress towards value. Right at the start, when defining ROI metrics can be hard, we can instead define conceptual models or themes for how we will assess ROI as projects progress. This helps make the case for investment in the project, and also ensures poorly performing projects are quickly stopped and sucessful ones reprioritised.

4. Implement & embed

Put systems in place to deliver your roadmap workstreams effectively.

Design systems to allow seamless data transfer across silos so everyone can see and use it. Create resources for accessing necessary data that ensure it is easily findable, accessible, and reusable.

Setup projects in ways which support collaborative working between data, IT and domain teams. Create shared spaces and make tools available for teams to communicate and share learnings.

But be pragmatic about what data is necessary. In some case, historic data may need to be migrated into a new database to deliver proposed projects. In others, it is new data streams that are most valuable. There is no need to spend ages getting data in order unless it is tied to delivering your use cases. But use the experience of bringing the data together to establish systems and practices that will make data capture easier in future and allow better quality data to build up over time.

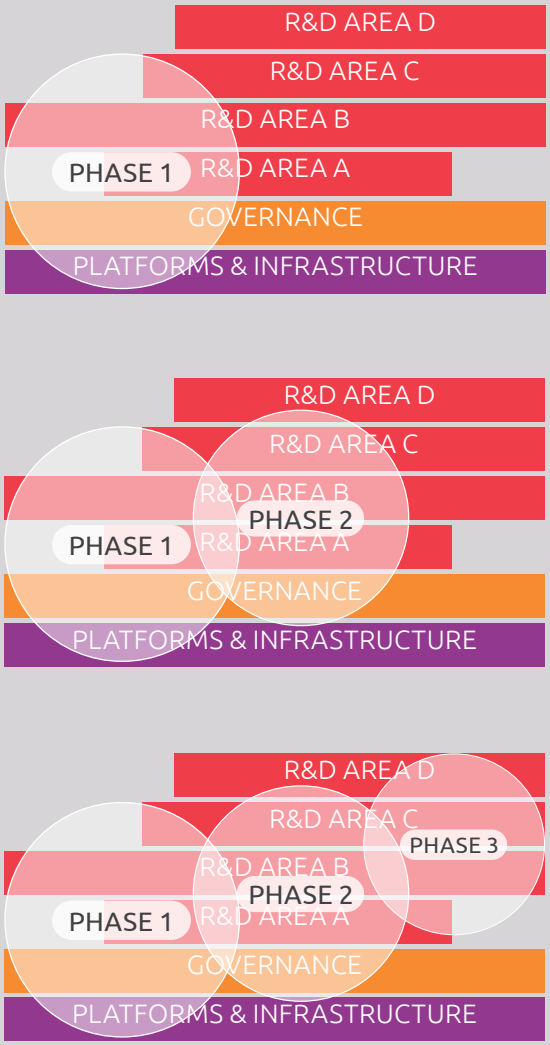
Finally, design projects so that learnings are taken on board. Encourage mentorship and places for teams to benefit from the expertise of others, or learn from outside consultants if they are driving this. Ensure everyone is learning as the projects progresses, so that teams can then take new skills and mindsets to future projects.

Bringing it all together to deliver a Digital R&D transformation programme

Whilst the above can be seen as four stages on the journey of an individual use case, taken together they underpin an integrated and evolving change programme. Often all four will run at the same time, and all should be learning from each other.

A Proof of Value will need some small-scale landscape work to understand which projects to take forward. But the complete landscape will be much more valuable and better understood if it is done with a view to informing a portfolio of use cases, which requires the Proof of Value to be completed.

Similarly, R&D digitalisation programmes will likely start with a rough roadmap, which will then be refined,



Trying to address every area of digital R&D at the same time is one of the most common ways that Digital R&D transformations can fail. Our approach is to understand the R&D landscape and the relative importance of each area. Through a series of targeted use cases that each bring real value to the business, we can address R&D areas in manageable phases while implementing key governance and infrastructure.

as use cases take shape and landscaping exercises are completed. This will then be constantly reviewed and evolved as new things are learned about the organisations data challenges, new use cases start, or successes present new opportunities. Implementation will be a continuous process.

And all of this is part of a continually evolving cycle. As projects advance, they will inform investments in technology and skills which will advance the landscape, and make it possible to start new use cases that were previously not possible.

This may sound complicated, but it is just another set of continuous process designed to increase the digital maturity of R&D departments – and so improve their speed and output. Any R&D department already has many such interlinked processes. Once you start it will soon become intuitive.

Resist the temptation to embark on a massive IT project to get all your data in order. Recognise that a more effective approach is to start with a few promising use cases, plan properly, and learn as you go, gradually using learnings to advance on your journey towards Digital R&D.

AUTHORS

Matt Jones
Lead AI and Data Science Strategist

David Hughes
Analytics Solutions Lead

Sam Genway
Senior AI Consultant

John Godfree
Head of Consulting

David Dungate
Senior Consultant

James Hinchliffe
Senior Consultant

Mark Knight
Account Manager: Consumer Industries

James Downing
Consumer Industries Strategy Lead

About Capgemini

Capgemini is a global leader in partnering with companies to transform and manage their business by harnessing the power of technology. The Group is guided everyday by its purpose of unleashing human energy through technology for an inclusive and sustainable future. It is a responsible and diverse organization of 270,000 team members in nearly 50 countries. With its strong 50 year heritage and deep industry expertise, Capgemini is trusted by its clients to address the entire breadth of their business needs, from strategy and design to operations, fuelled by the fast evolving and innovative world of cloud, data, AI, connectivity, software, digital engineering and platforms. The Group reported in 2020 global revenues of €16 billion.

Get the Future You Want | www.capgemini.com

For more details contact:

cprd@capgemini.com

MACS_GCS_20210812_MK