DataTechnology

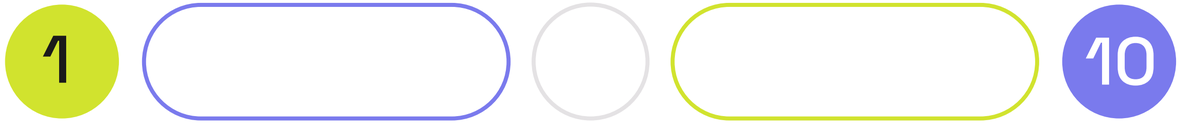
Cloud Data Platform Checklist

Contents

Cloud-based Data Platforms have significantly matured over the past few years. With the exponential growth in business data, many traditional on-premise solutions are outgrowing their capability, becoming less flexible, disjointed and costly.

Migration to cloud-base systems is affordable, flexible and completely scalable. However, without careful planning, a high percentage of data projects will fail. According to Oracle, more than 80% of migration projects either fail or exceed their budgets.[1]

These failures are completely avoidable. Based on our expertise in this area having successfully delivered a wide range of projects, we have created our 10-step checklist to help you plan and manage your Cloud Data Platform project.

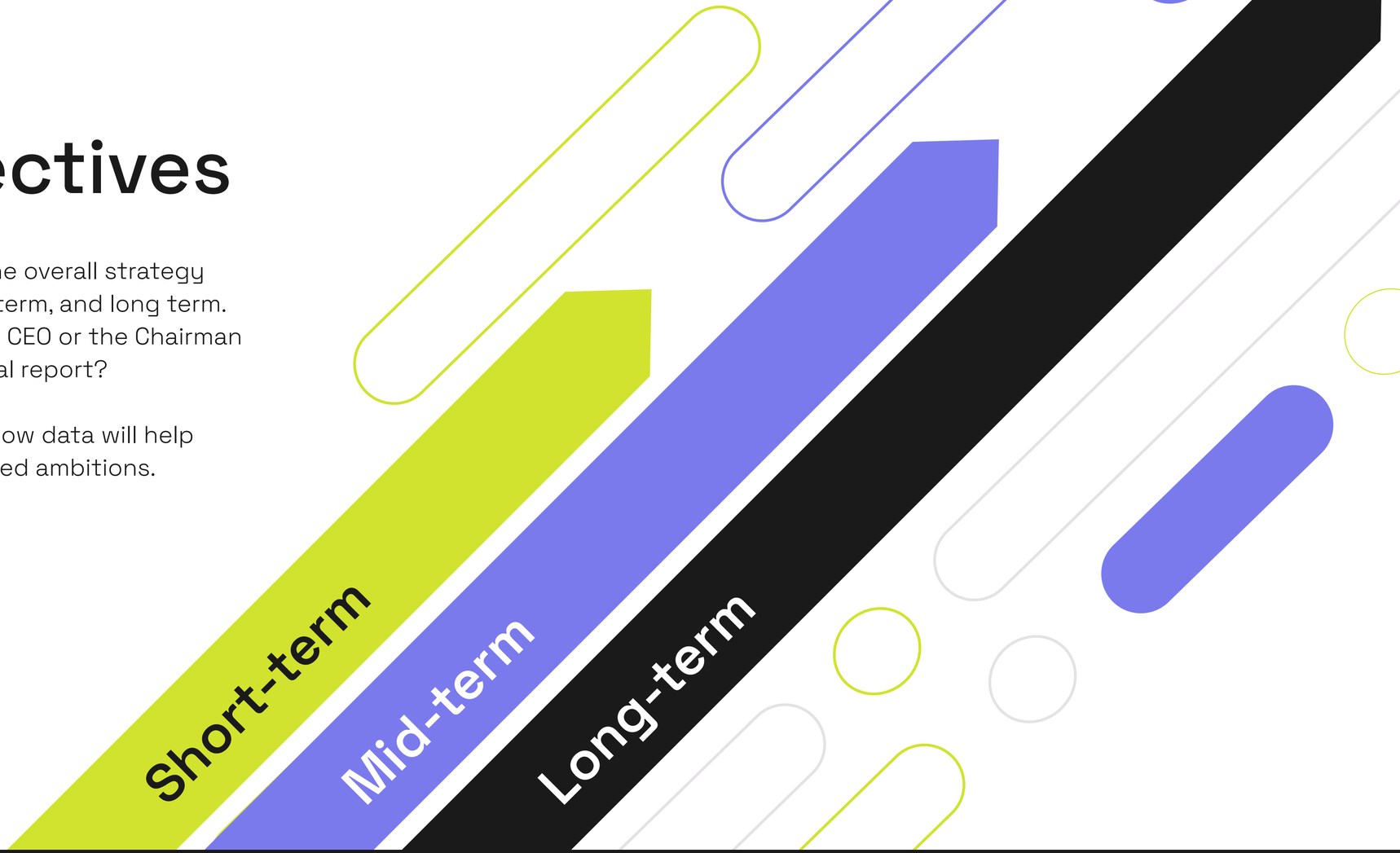


[1] https://www.oracle.com/a/ocom/docs/middleware/data-integration/data-migration-wp.pdf

1. Initial Objectives

The starting point is to look at the overall strategy of the business, short term, mid term, and long term. What goals have been set by the CEO or the Chairman and is this the focus in the annual report?

It is also important to establish how data will help the business achieve the identified ambitions. Document your main objectives.



2. Workshop

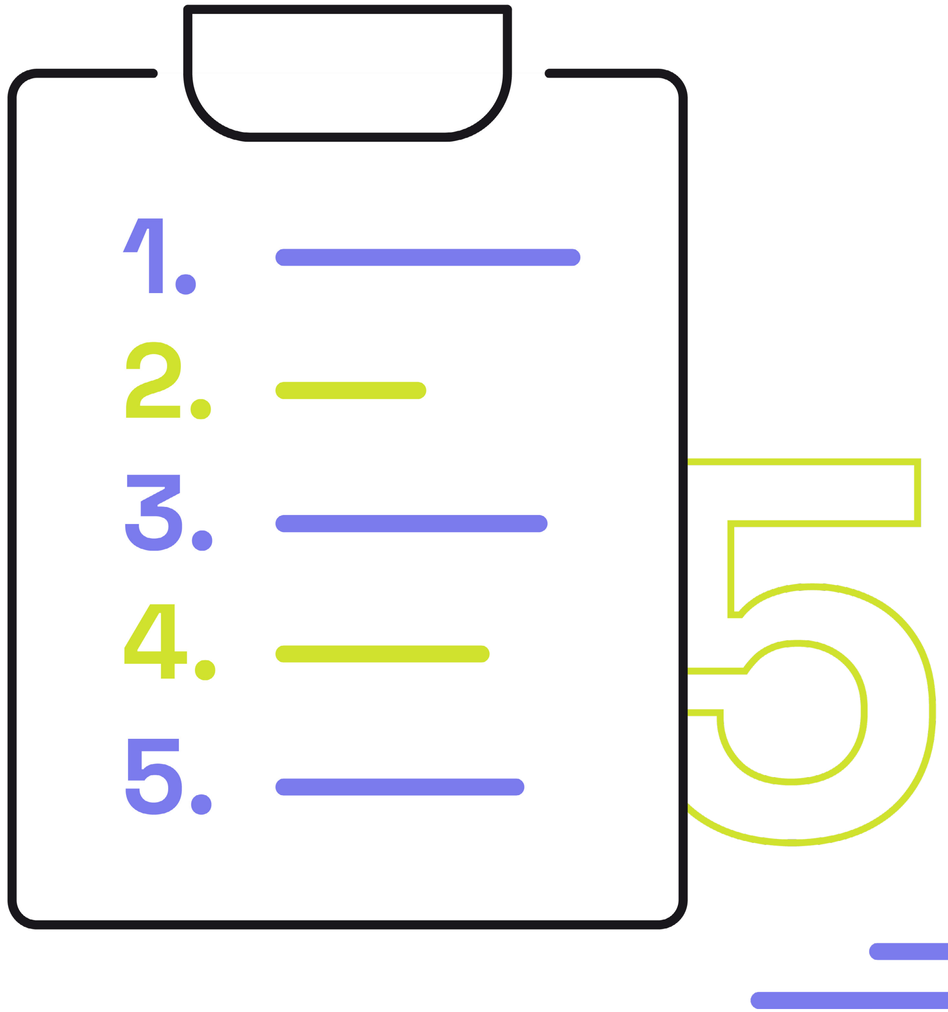
The next step is to start engaging with business stakeholders and plan a series of workshops. Sometimes referred to as Ideation Workshops, these are a great way to bring business stakeholders together and gain user buy-in which is a vital part of the process. This will facilitate department leads to present their use cases where they believe data will help them. Use cases might include things like speeding up the decision-making process, resolving a current issue, or making general improvements.

3. Use Cases

From the workshop sessions, shortlist the captured use cases to 5 or less. If you have a large number of use cases, then it is probably a good idea to group them by function or dataset.

Simply gaining traction with an idea is far more beneficial than stalling or procrastinating. Building a good use case also helps to get engagement by senior staff as well as project champions. On each use case, review KPIs and what good looks like. Eg. good service, good value, cost to serve. As well as senior stakeholders, also engaged with implementors within the business (those that make things happen... the doers).

Remember that on its own, data is just data, and not valuable per-se, until it has informed decision-making capability. Identify this with various stakeholders and build out a use case accordingly.



4. Consistency

Within large or complex organisations, there can be different metrics and classifications that define business processes – like different definitions of ‘cost to serve’. It is essential to standardise these so that they can be built off consistent datasets where everyone is agreed on the definitions. It is, therefore, good to have a robust data dictionary, that sets out a standard definition. This is also great for getting ‘user buy-in’ and engagement.

Search out users who can be project advocates, which will really help along the journey. When users see how you’re engaging with data that they are generating and managing, it can drive overall data quality, compliance and governance from their part.

Consider building a data quality App that can be used to highlight shortcomings in the current business data, and be a catalyst for providing additional support and training resources needed within the business.

5. Proof of Value

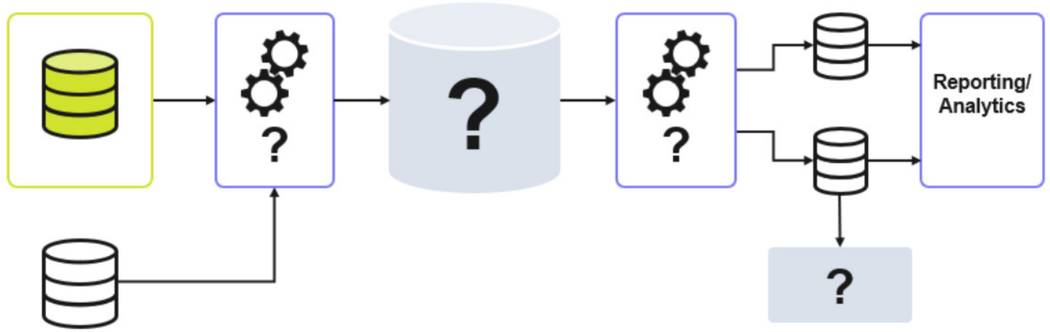
It is better to ‘fail cheap’ than risk any large commitments. Therefore start small and take an iterative approach by building an initial proof of value. It is easier to deliver this than trying to scope, define and build the ‘mothership’. Strive to produce a self-contained proof of value that can be assessed on its own merits.

Get people excited about data. Engage them. Involve them at the beginning of the design process. What KPIs are going to help you? The more you get them inspired in the beginning, the better user adoption and engagement you will have throughout the project.

A proof of value needs to have substance and demonstrate real benefits within the business.

Data literacy is a key element and should be considered in the early stages. You want users to be comfortable with their dashboards, and be able to use the information at their disposal. You want them to ask questions, engage with content and be able to make good insights. The proof of value is the best time to identify knowledge and data literacy gaps and put a training plan in place.

6. High Level Architecture



Before you can answer this, establish who your users are, and what they will want from the platform. Will the data need to be live or batch processed every day? Perhaps something in between? Think about the operational practicalities. For example, international firms will operate across several time zones. What does the 'end of day' look like for them? When is the data refreshed?

Is the source data structured, unstructured or both? This may channel you down a particular solution and look at a Data Warehouse, a Data Lake or a hybrid approach.

Cost will play a large part in the overall solution, so look at the different cost models. Is there a cost per GB for storage, what about inbound and outbound data bandwidth?

Do you want to code your own data pipelines using a particular coding language, or are you looking for no code tools like Azure Data Factory, Qlik QDI, Arcion or Hevo?

There is also much overlap between the big cloud providers and there may be an argument for using the same provider as your existing technology stack, for example Microsoft, Google or another.

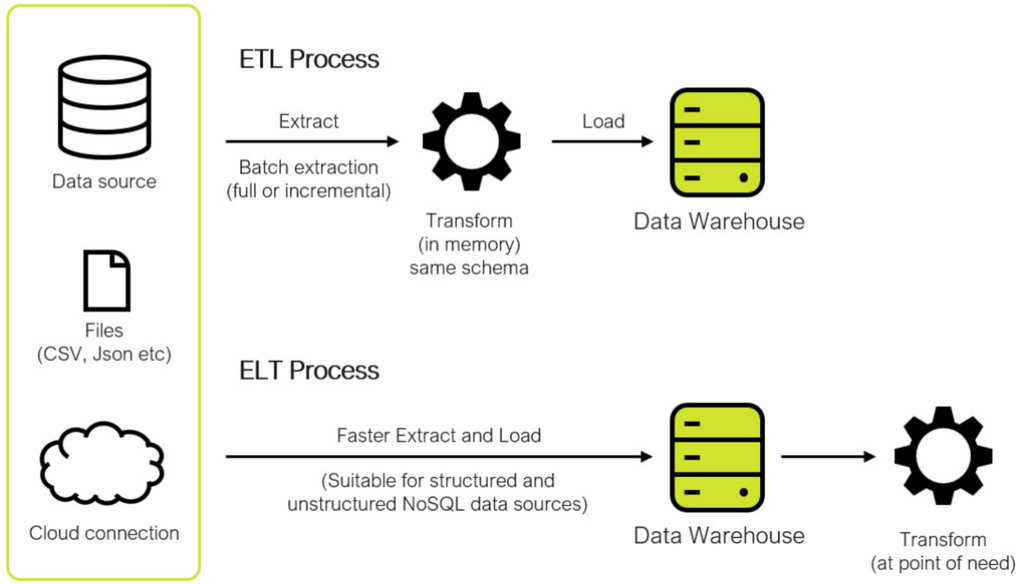
Microsoft Azure

Google Cloud

aws

Pick a product that has continuous development and that will evolve and mature, based on its past performance. Keep aware of developments in the market and the sentiment from users. Be careful of using niche products and make sure you do your homework on them.

Avoid being an early adopter or using any brand new offerings that may inevitably contain bugs and issues. It's always better to use tried and tested solutions and software versions that are one behind the current iteration. Review case studies of the particular product in action.



The traditional data pipeline centres around an ETL process which extracts data, then transforms it before loading it into a Data Warehouse.

ETL pipelines are usually ridged and require a consistent schema for the transform process. Data ingestion is usually done in batch mode, either extracting the entire data source or an incrementally changed subset of data. ETL pipelines typically don't work well with unstructured data or real-time data feeds.

In contrast, ELT pipelines are concerned with extracting and loading source data into the Data Warehouse before it is transformed for the reporting requirements. Since the transformation is done after data ingestion, ELT pipelines can work with structured and unstructured data and are generally faster at loading information. Transformed data can be reloaded into the Data Warehouse and consumed by reports at the point of need.

Cloud Data Warehouses tend to include the flexibility to work with either pipeline, or a hybrid of the two.

Do you load data from source (Bronze standard) or undertake basic transformations or validation (Silver Standard) and where is your analytics-ready data (Gold standard)? Thes are questions that need answering.

7. External Support

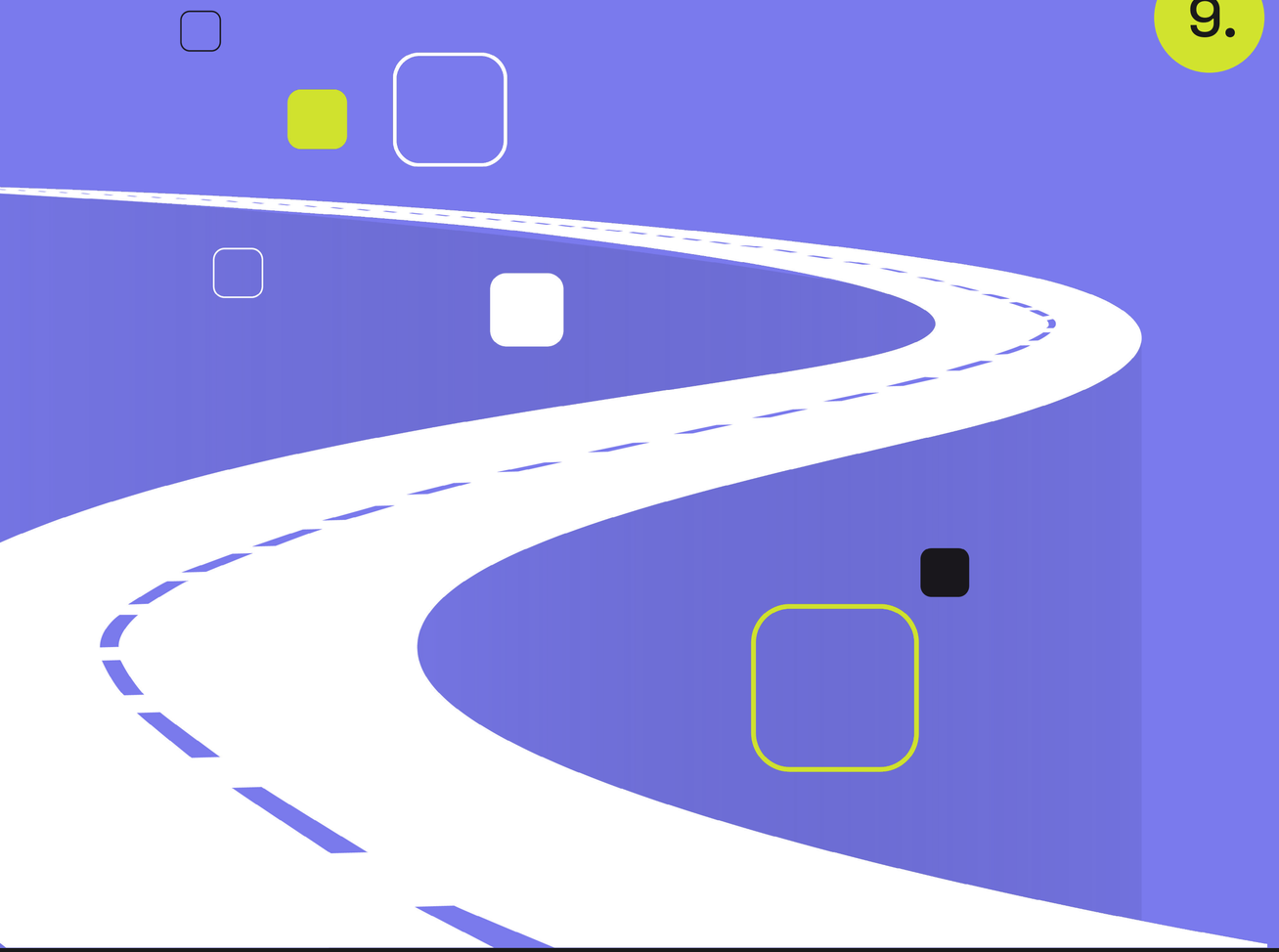
You may have many of the required skillsets inhouse, but engaging with an external partner with expertise in the field will allow you to validate your approach. A blended approach will allow you to bring in the right resources for the project, at the right time. You will also get an objective view.

Look longer term and consider the post project view. Ensure you have the capability and resources to successfully support and run the platform after it has been delivered. Requirements will evolve over time so it is important that you can manage the future.

8. Data Security and Compliance

In the early stages, plan who will need to access the data, for what purposes and at what level. Design a Data Security and Compliance policy. Is your data protected? Do consumers only have access to relevant content, are they authorised to use it?

Can you successfully ensure that regional data is only accessed by the correct regional team? What about customer data? Can account managers just see their own clients, accessing specific tariffs and rates?



9. Future Roadmap

Keep setting short term, mid term, and long term goals.

Create usage reports to see which Apps are popular and generate value within the business. Don't be afraid to remove anything that has gone stale – it is better to have valuable insights rather than too many options. Keeping the platform relevant and healthy is the key to longevity.

From time to time, bring in outside consultants who can provide an objective viewpoint of your platform, to help establish areas of improvement as well as validate which sections currently work well.

10. Continuous Improvement

Keep the platform in good shape as it grows.

Undertake regular health checks. Review and optimise your platforms. By their nature, cloud platforms are constantly evolving, with new services coming on board at regular periods. Current features will inevitably be ‘sunsetted’, other features will evolve and mature. There may be syntax changes to coding languages, or updated user interface elements from time to time – not always immediately beneficial.

Take time to step back and review what you’ve built. Is it still providing value and relevance? Is it still answering the business questions that people are asking?

As stakeholders continue to utilise visualisations, they will undoubtedly want more capability and better insights over time. Rather than just provide a ‘what’s happening now’ snapshot, consider measures over a time period for ‘like for like’ comparisons. Has a particular marketing campaign had an impact on the figures?

Business requirements will undoubtedly change, so it is important that the Business Intelligence teams do regular reviews with stakeholders to identify which elements are working well, and which need improvements. New tools and feature sets will provide new possibilities, and allow you to demonstrate future insights that are not available today.

For example, a lot of businesses are committed to their Environmental, Social and Governance (ESG) frameworks. Future ESG requirements and standards will no doubt bring questions that mature data insights can help answer.

DataTechnology

At Data Technology, we help businesses migrate their data onto the major hyperscaler platforms, such as Microsoft, AWS and Google, Snowflake? Where it becomes connected, readily available and analytics-ready.

By taking a systems-agnostic approach, we marry business problems to the right technical solutions, and then pair those technical solutions with the right toolsets. Our tailored implementation and managed support services help to make sure our clients benefit from the convenience, security and reliability afforded by cloud solutions.

Get in touch to discuss how Data Technology can support you with any part of the data migration journey.

Call us
0845 6588 480

Email us
info@datatechnology.co.uk