

AHEAD

AHEAD Playbook

Innovation Through Modern Data Practices

Modern Data Is Changing the Game

In the modern data environment, we're finding ourselves with a wider set of users, new ways of working, and new expectations for what our data can tell us. Users want broad access that will allow them to explore data in ways that we've never needed to support in the past—a requirement that can put tremendous strain on enterprises as they navigate data model degradation, the demand for more storage and computing power, and an ever-increasing list of security concerns.

In this playbook, we'll explore two facets of modern data:

- [Choosing the Right Environment for Analytics & Data Science](#)
- [How Agile Thinking Can Fuel Innovation in Data Analytics](#)

Each section within this playbook provides insights into the best practices for modern data analytics and the many benefits to the adoption of these methodologies. For a deeper dive into each topic, watch the on-demand sessions from our complementary [Accelerating Innovation Through Modern Data Practices Summit](#).

Choosing the Right Environment for Analytics & Data Science

So, how do we satisfy the new and still-emerging demands of modern data? In the traditional model for managing data analytics environments, we've used local or personal machines and sparse integrations on a common set of servers or mimicked the software development paradigm with a dev-test-production environment. However, neither of these methods are conducive to meeting the requirements of modern data. In fact, according to Dataiku, as much as **75% of the data in traditional environments is completely unused**—in what is often referred to as the “data swamp.”

Given that we need to become more exploratory with our data, more fastidious with our security, and more deliberate in our approach, where should organizations focus their energy to fulfill the needs of modern data?

1. Determining how to adopt a fail-fast approach
2. Identifying where to insert governance
3. Supporting variability in computing and storage requirements
4. Supporting the need for on-demand data expertise
5. Securing the environment with the right level of access

Modern Data Calls for a Dynamic Analytics Environment (DAE)

By utilizing a range of different technology enablers and solution components, we can solve for many of the demands of modern data by creating what is referred to as a Dynamic Analytics Environment (DAE). This allows us to change our environment on a case-by-case basis to mitigate the challenges of both variability and demand, leaving behind the one-size-fits-all approaches of the past.

Requirements

- Embrace fail-fast
- Provide quick, secure access to data
- Knowledge sharing
- Ability to scale environment on demand
- Ability to govern and monitor use, cost, and access

Technology Enablers

- **Cloud** - Cloud resources to provide elasticity and access to new services
- **IaC** - Build environments automatically
- **Large Scale Storage** - Data Lake/Lakehouse/Mesh to centralize data in a data hub
- **Data/Compute Separation** - Snowflake/Starburst/Presto/Trino

Solution Components

- Separate work into “Cases”
- Define environment “templates” for different classes of problem
- Deploy “Case Environments” on demand based on environment templates
- Pre-validate environment templates with security team once, use many times
- Centralize structured and unstructured data

Benefits of DAE



Case
Compartmentalization



Clear
Governance Process



Fail-Fast
Support



Resource
Reclamation



Elastic Resource
Consumption

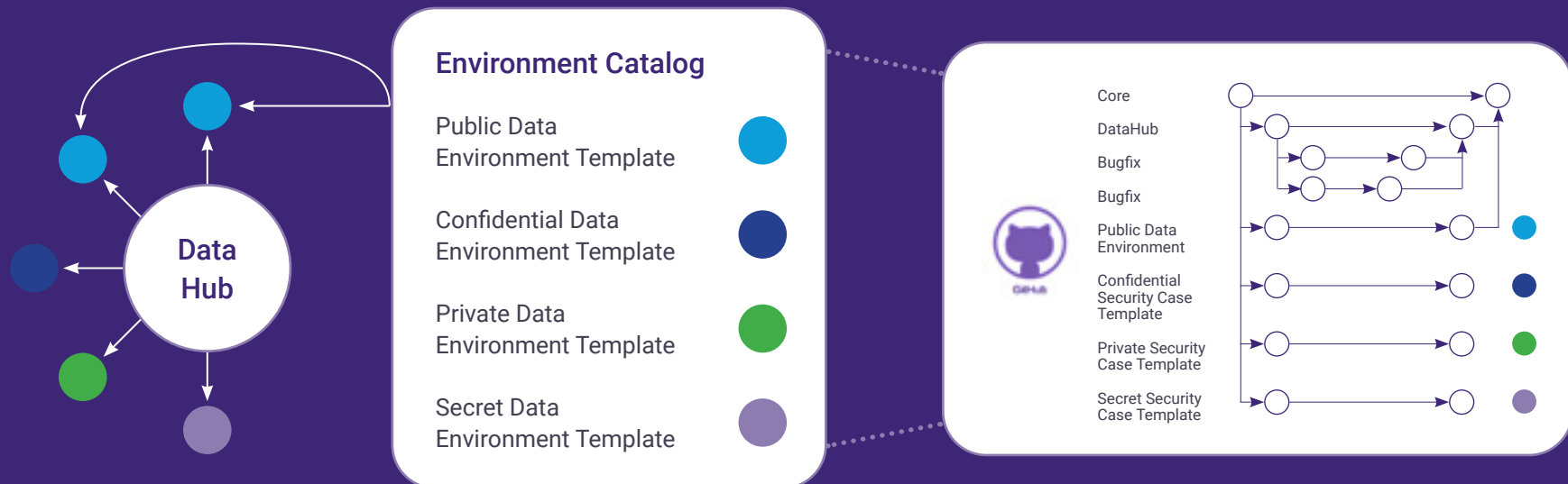


Pre-Authorized
Security

DAE: Conceptual Architecture

Environment templates can (and should) come in a variety of flavors. We can facilitate this by using a central data hub to hydrate specific environment templates per case or workflow.

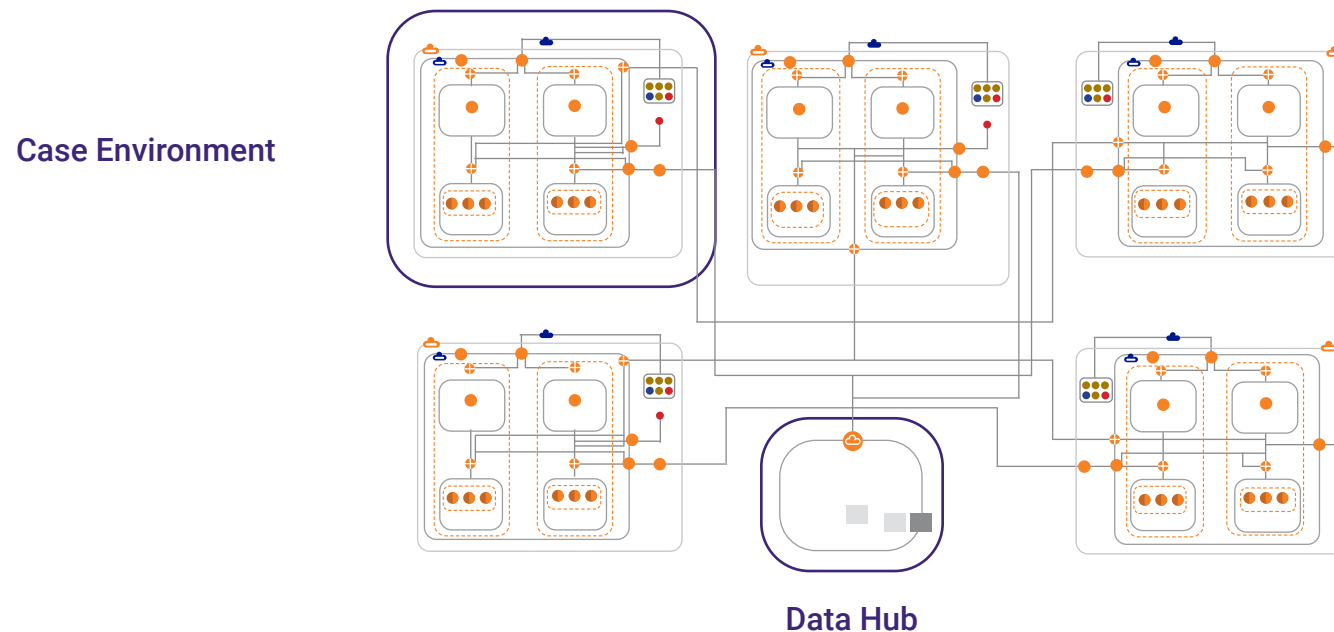
In the example below, the environment templates are organized by access level—from public to secret data—in order to provide users with only what they need or what they are authorized to access. The code to build these individual environments is also source-controlled, giving the organization more control and visibility over what is being built and how it is being supported.



DAE: Environment Services

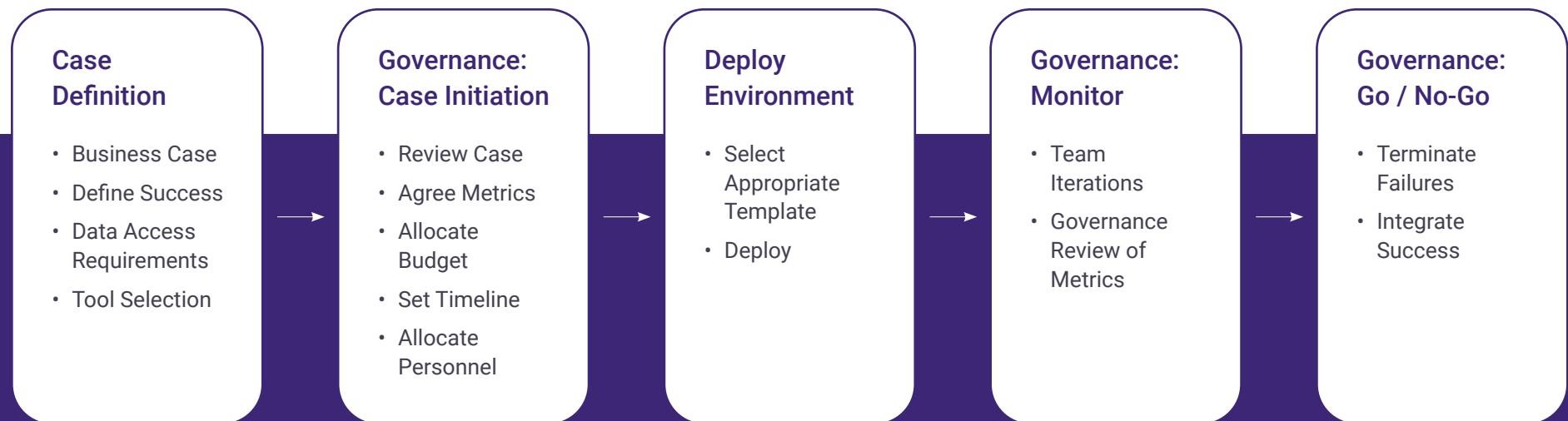
What would implementation of a DAE look like? In the example below, you can see how the data hub feeds into specific case environments. These environments are, in this case, self-contained virtual networks with all of the services that are required by the case template, such as a development environment, a cluster compute engine, and local data storage.

Since the environment is self-contained, all of the security tools and monitoring can be built in just once. Data can be fed into each case environment using written scripts or by giving the environment direct access to the data hub—the answer to which is best is usually determined by the amount of data that needs to be moved.



DAE: Governance

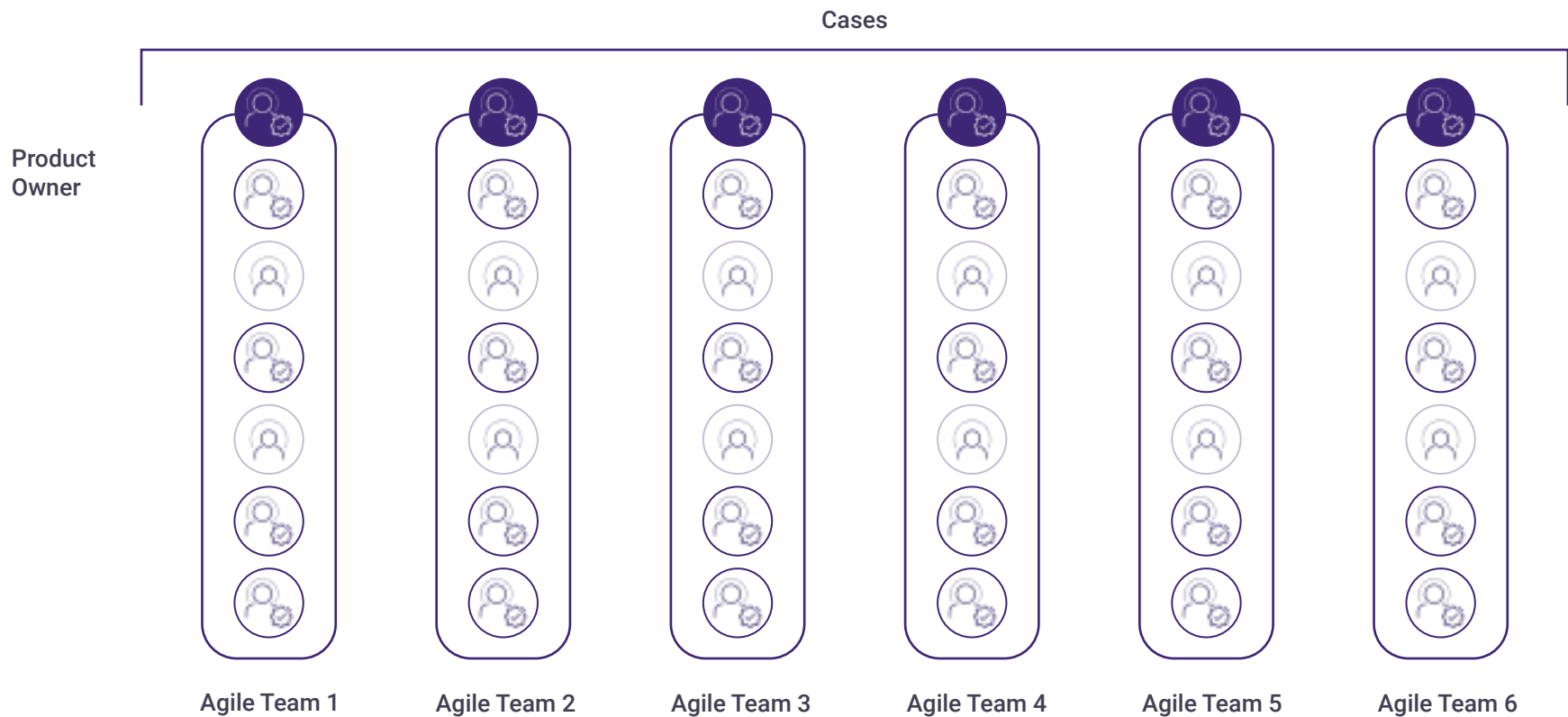
Beyond elasticity and security, DAE provides a useful mechanism for inserting governance and control. Before environments are hydrated from the data hub, key case management metrics—such as a definition of the case itself, a definition of success, data access requirements, and the necessary tools—can be added to each case environment. This enables a fail-fast approach, provides clean separation between cases, clarifies cost and implementation details, and rids us of the aforementioned “data swamp.”



DAE: Team Structure

What do team structures look like? For starters, they should be long running, with the team owning the discovery, development, and ongoing operations of the analytics.

Teams are best organized using agile thinking with a cross-functional makeup—product owners, data analytics experts, technical experts, and visualization experts are all integral components of any team working within a dynamic analytics environment.



How Agile Thinking Can Fuel Innovation in Data Analytics

In virtually any organization today, the ability to produce high-quality digital products faster has become an imperative. Accelerated innovation is no longer an option for enterprises that wish to remain competitive and thrive in the digital economy—and data analytics is no exception.

Defining Data Analytics



Science of analyzing data to find conclusions that help organizations optimize performance and make business decisions



Involves the transformation of raw data into digestible formats fit for human consumption



Deliverables typically include digital tools, reports, and dashboards

Realities Facing Data Analytics Teams



Data is usually tied to products, with different technical expertise required to leverage dynamically (APIs, languages, etc.)



Not every insight requested is a priority, and not every insight is needed long term



Organizations are frequently subject to regulations and laws related to their data (SOX, HIPPA, etc.)

Data Analytics: Where Are We Today?

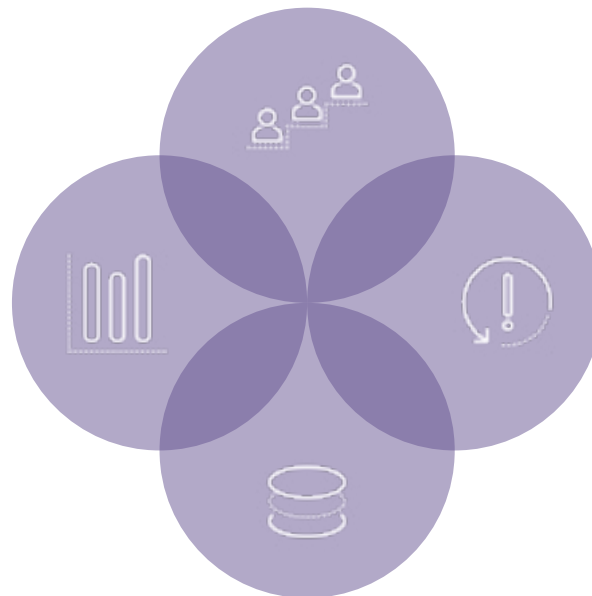
Given the realities currently facing data teams, organizations frequently adopt a fairly outdated approach:

- Insights are built as projects using temporary or non-static teams
- Teams are kept in silos and given full governance over data
- Traditional project management methodologies are used under the assumption that the end consumer knows what they need and that priorities will never change

And with these “old-school” approaches to data analytics come a host of problems and inefficiencies:

The first delivery of insights is very late in the project, after a large amount of time has passed and the insight is likely no longer relevant or accurate.

Product owners do not have governance over the data portion of their product stack.



The original team that built the insight is no longer intact when bugs need to be fixed, new insights added, and refreshed data sources included.

Large quantities of interdependencies with product teams are created and need to be frequently handled and resolved.

Solve Data Analytics Inefficiencies Using Agile

Agile can be a useful way to solve the problems introduced by outdated methods around data analytics, as it serves to enhance the “people” and “process” components of any given project. Combined with the right tools and technologies, agile thinking can alleviate most—if not all—problems related to data analytics team inefficiency.

In order to embrace agile, however, an organization must follow 12 core principles across four core tenants:



Delivery

- The most important priority is early and continuous delivery of working technologies that solve problems for end-users
- The shorter the frequency of deliveries, the better
- Working technologies are the primary measure of progress



Collaboration

- Individuals should be motivated, supported, and trusted to get the job done—not micromanaged
- Business and technology people must work together daily
- Collaboration must be sustainable and continue indefinitely



Teamwork

- The best technologies are designed, built, and maintained by cross-functional, self-organizing teams
- A face-to-face conversation is the most efficient and effective method of communication
- Teams should regularly reflect on how to become more effective

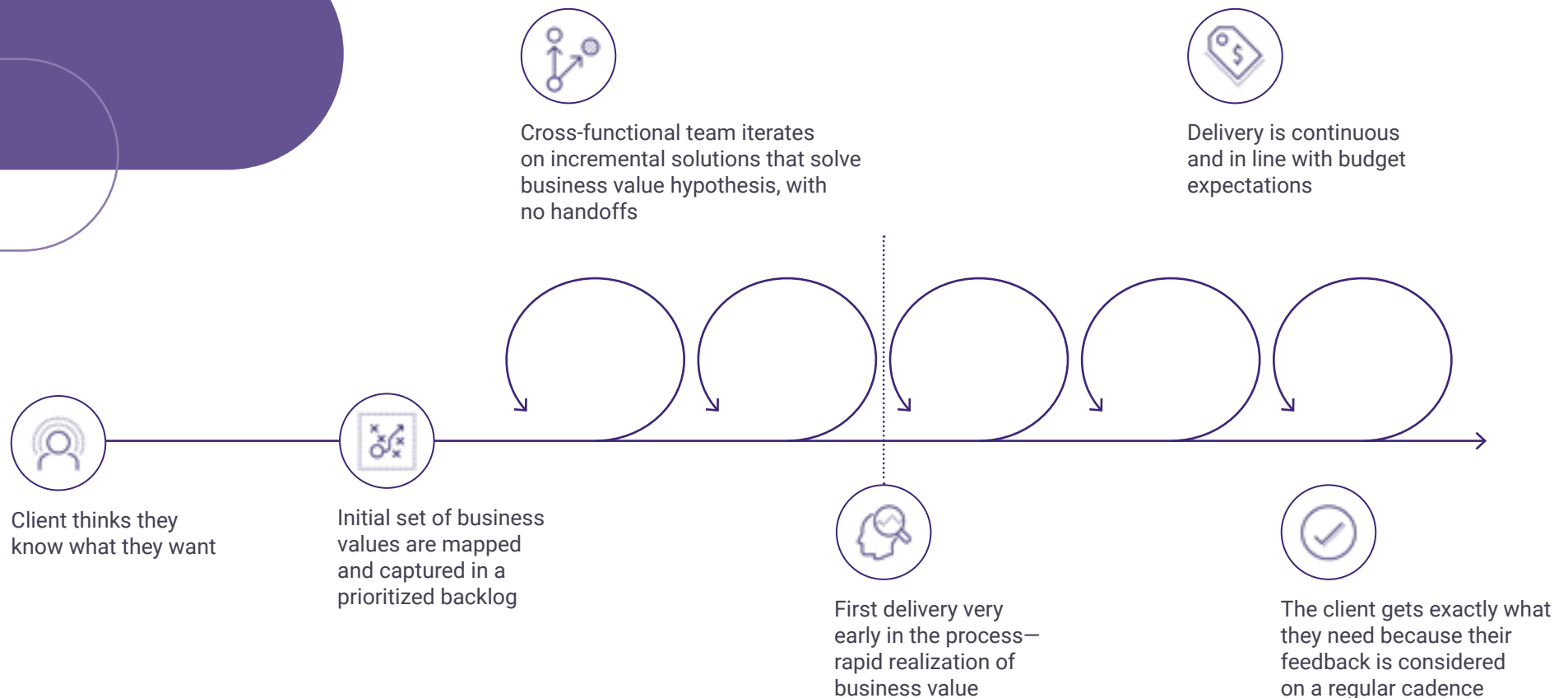


Excellence

- Changes to requirements are always welcome
- Continuous attention should be given to technical excellence and good design
- Simplicity—the art of maximizing the work not done—is essential

How Agile Works

Agile thinking allows teams to tackle projects using an iterative methodology, providing clients with a rapid realization of business value in best-case scenarios and enabling teams to fail fast and pivot toward a new strategy in worst cases. This gives teams the ability to validate their approaches and goals before a project has been completed, therefore maximizing overall efficiency.



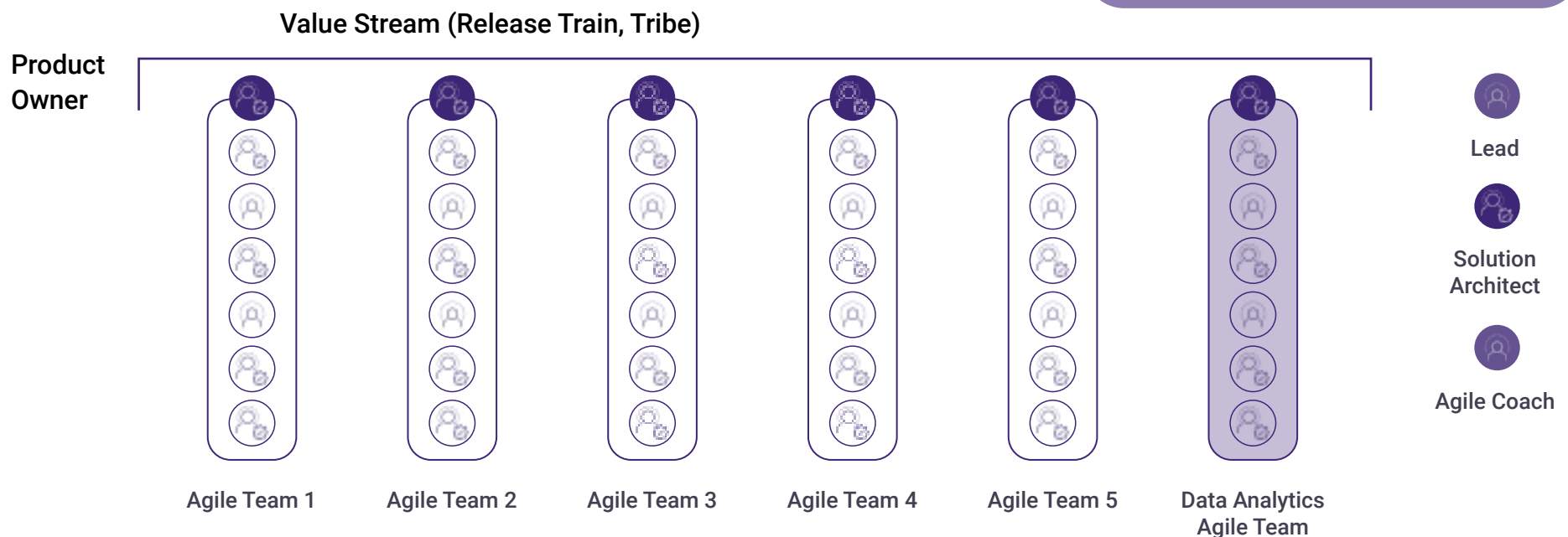
How Does Agile Fit Into Data Analytics?

Historically, data analytics teams have been kept siloed from product delivery teams, ultimately disconnecting them from any specific value stream. To break these silos, organizations have adopted an approach where data resides with the product teams. But this approach also presents challenges such as heightened data security risk and interdependencies as well as a lack of governance, data cleanliness, and product context.

An agile methodology serves to integrate the data experts into the value stream by creating a data analytics agile team, complete with its own product owner.

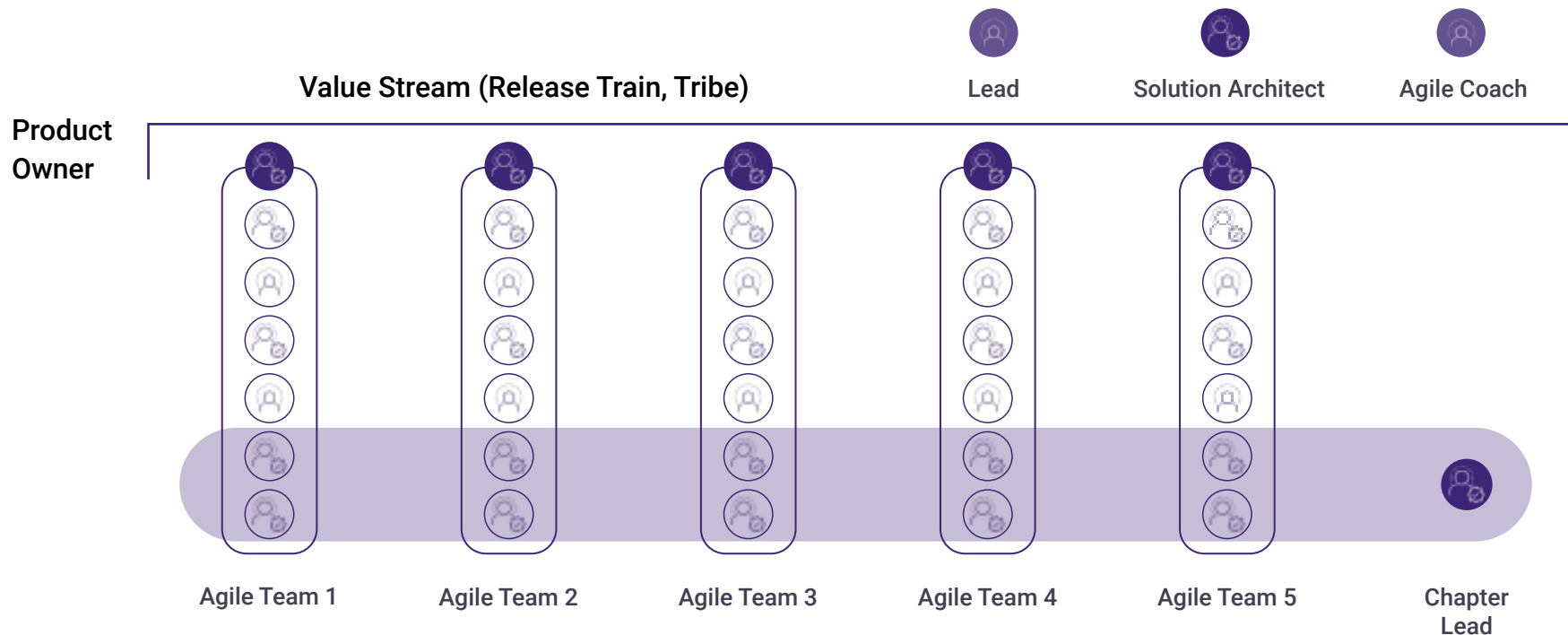
This approach can ultimately create another silo within the value stream itself, and tends to work only when:

- Data residency is common across all products
- No other value streams leverage data
- Data analytics is considered an offering



A Better Approach: Embed Data Analytics Expertise

The best way to leverage agile thinking to improve data analytics is to create a data chapter within each agile team. With this approach, data analytics experts are embedded into each workflow, allowing them to verify that the data analytics solutions are being met across all teams. This enables end-to-end delivery, eliminates interdependencies, provides product context, and streamlines data governance, all of which serves to not only enhance the efficiency of data analytics teams, but solve problems related to data dependencies across the entire value stream.



Key Takeaways



AHEAD

About AHEAD

AHEAD's Data Analytics expertise helps enterprises leverage their data to fuel innovation and deliver better outcomes for customers. For a deeper dive into the topics discussed in this playbook, watch our complementary [Accelerating Innovation Through Modern Data Practices Summit](#).

To start the conversation with one of our Data Analytics experts, [get in touch with us today](#).