

Statement of Architecture Work

Data Platform project  
Vår Energi 2020

Table of Contents

[1 Statement of Architecture Work 4](#_Toc44508239)

[2 Objectives and Scope 8](#_Toc44508240)

[3 Roles and Responsibilities 10](#_Toc44508241)

[4 Architecture Approach 12](#_Toc44508242)

[5 Work Plan 15](#_Toc44508243)

[6 Risks and Mitigations 20](#_Toc44508244)

[7 Acceptance Criteria and Procedures 21](#_Toc44508245)

[8 Signature Approvals 22](#_Toc44508246)

Document Information

|  |  |  |  |
| --- | --- | --- | --- |
| **Project Name:** | Data Platform project | | |
| **Prepared By:** | Renate Endal, Christine Toft Stenersen, Johan Ludvig Brattås and Morten Taksrud | **Document Version No:** | 1 |
| **Title:** | Statement of Architecture Work – data platform project | **Document Version Date:** | 01.07.20 |
| **Reviewed By:** |  | **Review Date:** |  |

Distribution List

| **From** | **Date** | **Phone/Fax/Email** |
| --- | --- | --- |
|  |  |  |
|  |  |  |

| **To** | **Action\*** | **Due Date** | **Phone/Fax/Email** |
| --- | --- | --- | --- |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

\* Action Types: Approve, Review, Inform, File, Action Required, Attend Meeting, Other (please specify)

Document Version History

| **Version Number** | **Version Date** | **Revised By** | **Description** | **Filename** |
| --- | --- | --- | --- | --- |
|  |  |  |  |  |
|  |  |  |  |  |

# Statement of Architecture Work

## Project Request and Background

Vår Energi is one of the largest E&P companies on the NCS and have an ambition to grow. In 2019 Vår Energi established a Digital Transformation Program, as it was expected from the owners, the competitors worked on digitalization and employees were eager for new ways of working. When defining their digital vision and roadmap, various digital initiatives were defined to make sure Vår Energi reach their digital vision of “turning bytes to barrels”.

Vår Energis **current IT landscape lacks a common way of accessing, sharing and collating information across units and systems**. This makes it difficult to do analyses, make reports, build dashboards that involves information from across several IT systems and business units, with the current IT landscape. In addition, it is nearly impossible to collate information from the multitude of systems as there is no common way of modeling the information found in these systems. It requires a large amount of manual work when collating and consolidating information, making the processes inefficient and time consuming.

A common data platform will be able to ingest the information found in the different IT systems across the different business units and prepare it for use. It will thus be an **essential enabler for many of Vår Energi’s planned digitalization initiatives** and should extend the capabilities of the current underlying technology platform, information availability and governance to support Vår Energi’s digitalization roadmap and strategy. Prioritized initiatives that have the Data Platform as a prerequisite for success is Energy Management System, Production Dashboarding, Mobility in Field Operations and Supply Chain Control Tower. Additionally, there are several initiatives in the pipeline, such as Monitoring of Heat Exchangers, Mechanical Integrity Management, emulsion management, integrity chemicals

In addition to enabling the digital initiatives **the data platform will also enable ICT to become more efficient**. ICT currently manages an IT landscape without any integration platform, which makes it difficult to ensure data flows between different systems and applications. The data platform will provide a mechanism that can make ICT more efficient and effective in their activities, by easier configuration and management of data flows.

## Project Description and Scope

The project has followed the established project set-up in Digitalization. Thus, the Data Platform project has been divided into two phases: Design phase and Implementation phase. The design phase set the foundation for the implementation phase, with a suggested architecture for the implementation. This foundation will be further elaborated and detailed in the current implementation phase. Therefore, both project phases are included in this document.

**Designing the Architecture - Design Phase**

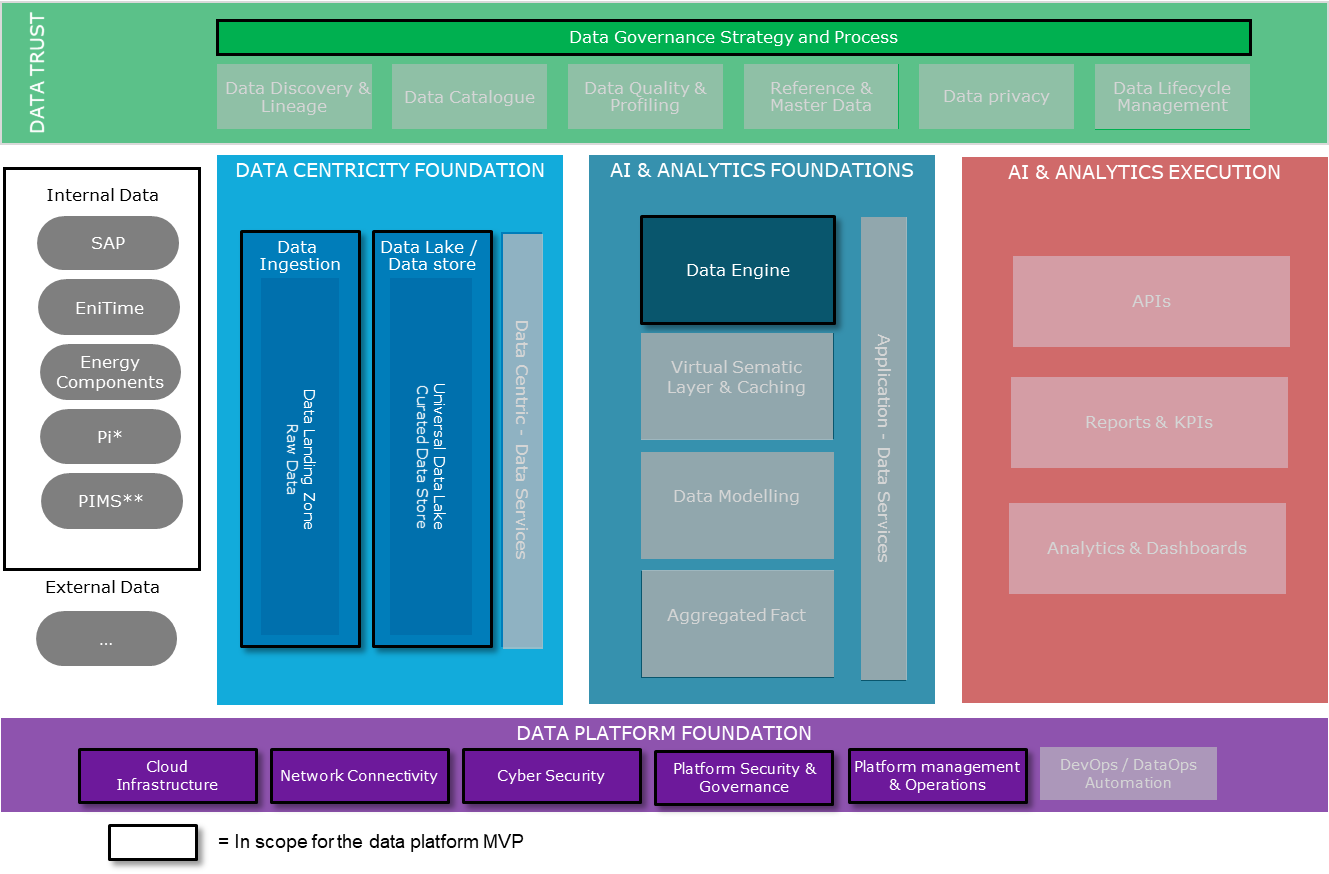
The objective of the **Data Platform Architecture** project was to propose an architecture, technology recommendations and suggested next steps for a data platform for Vår Energi

Thus, the scope for the Data Platform Design Phase was to Design the Data Platform Architecture and evaluate products to acquire, with the objective to lay the foundation for implementation of the Data Platform. The foundation of the data platform will be a minimum viable product (MVP) and the platform will be further developed in iterations afterwards.

**Detailing the Architecture - Implementation Phase**

The scope for the Data Platform Implementation Phase is to implement a Common Data Platform MVP in Vår Energi. The scope for the Data platform MVP is determined based on the initial recommendation for the whole data platform in the design phase, the chosen case that we are setting up and what we can deliver within the budget and timeframe of the project.

The highlighted components in the architecture drawing are included in the MVP scope:

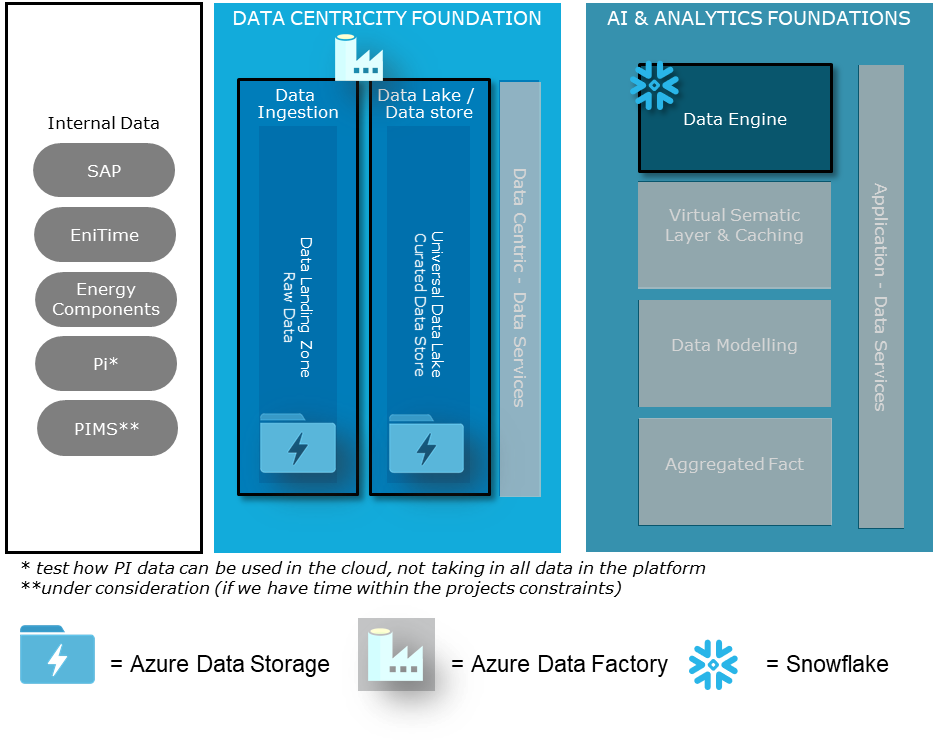


**Scope for Data Platform Foundation**

* Set up an **Azure subscription** for Vår Energi
* Set up the required **Azure infrastructure** to fulfill the data platforms requirements, but also taking into account that the Azure platform will be used by other projects/initiatives/workloads in the future and prepare for that
* Establish **a minimum of governance and security measurements** such as create a naming standard and a minimum of policies to be enforced in the platform
* Do a **cyber security evaluation of the Azure platform** and get it approved. This will apply for the whole platform so that it won’t be needed to do it again for future initiatives that would like to use the platform
* **Ensure platform management** of the Azure platform. Atea will be responsible for this and the project covers the costs throughout 2020

**Scope for Data Sources, Data Centricity Foundation and AI & Analytics Foundation**

* Establish **connections between the data platform and the selected data** **sources** using Azure Data Factory (ADF) and integration runtimes (on-prem agents for ADF)
* **Build ADF pipelines** for copying data from sources to raw data lake. (Scheduled batch jobs)
* We will test **how to copy PI data to cloud** 
  + PIMS will be included if we have time
* Build **transformation pipelines in Data Factory for SAP** data and load data into curated data lake
* Set up and **configure Snowflake**
* Create ADF pipelines to **copy data into Snowflake**



**Scope for Data Trust**

We will focus on Governance Strategy and Processes within Data Trust and have deliverables within Data Lifecycle Management and Data Platform Governance

***Data Lifecycle Management deliverables:***

* **Naming conventions:** A holistic standard for labeling data objects that helps to ensure a “common language” for data objects in the organization, which helps to keep work organized and understandable
* **Data Quality:** Establish data quality parameters from the Data Platform perspective and create rules for validation and consistency check of data coming from the source systems.
* **Structure and organization of data:** As part of the Data Platform Model we will look at how the data should be structured and organized and create rules for storage organization to increase usability and simplify security
* **Data lifecycle management:** A set of policies for managing the flow of data throughout its lifecycle and defines the steps from collection and initial storage to archiving and purging

***Data Platform Governance deliverables:***

* **Development and technical guidelines:** The guidelines apply to software developers, data engineers and architect, and represent a set of rules / standards for implementation of new features and modification of existing one
* **Maintenance and support processes:** Procedures for maintenance and support of the Data Platform aligned with the existing processes in Vår Energi
* **Request/approval/implementation for new reports and data sources:** Guidelines for request, approval and implementation process for new data sources, which will establish common routines for extension and ensure that the Data Platform is not overloaded
* **Define roles and responsibilities:** Adapt the DP to Vår Energi IT operations with defined roles and responsibilities, and define the organizational structure to support the Data Platform, incl. service and maintenance

## Strategic Alignment

See Chapter 1.1

# Objectives and Scope

## Objectives

The business objectives of this architecture work are as follows:

* Flexibility to start small and scale as needed and utilize services and software that can scale without license-changes according to usage
* Integration towards 3rd party vendors to get best solution to support both current and future business needs while retaining independency and minimize vendor lock-in
* Enable the use of many different end-user tools and applications to best fit end-users’ purpose and thus increase efficiency
* One single-source-of-truth
* Make it easier to find data through data and system documentation
* Increased data sharing, reuse and delivery across the business functions which will enable more efficient business operations by breaking organizational silos

The business objectives can be further specified when the it’s been decided on how to further develop the data platform and which initiatives that will begin to use the platform.

## Scope

See Chapter 1.3

## Stakeholders, Concerns, and Views

|  |  |  |
| --- | --- | --- |
| **Role** | **Concerns about the business** | **Concern about the Data Platform** |
| Digitalization Manager | Responsible for the Digital Portfolio and concerned about implementing digital initiatives in the organization | Data Platform important enabler for several digital initiatives and the digital portfolio, and the digitalization manager is thus concerned for the progress of the project as it impacts the progress of the digital program |
| ICT Manager | Concerned about making sure that the ICT department delivers on the business unit’s needs. | Project owner for the Data platform implementation project. Concerned about implementation progression, budget and the quality of the delivery. |
| Architecture Board |  | How the Data Platform Architecture affects and aligns with the existing architecture and the organization’s needs. |
| Infrastructure Team | Ensuring that the infrastructure support the business needs/requirements | Concerned about security and efficiency of the infrastructure in the data platform and how it integrates with the rest of the IT architecture in VE. Additionally, they are concerned about the management of the infrastructure. |
| IM Team | Enable the business to use data in an efficient manner | The Data Platform is intended to store and manage a lot of VEs data. IM team is concerned about data lifecycle management in the data platform. |
| Digitalization Team | Enable digital transformation in the organization | Data Platform important enabler for several digital initiatives and the digital portfolio, and the digitalization manager is thus concerned for the progress of the project as it impacts the progress of the digital program |
| AM Team | Ensuring that the applications support the business needs/requirements | The AM team is responsible for managing the applications in Vår Energi. They are concerned about making sure that the applications in the data platform are managed in a secure and efficient way. |

# Roles and Responsibilities

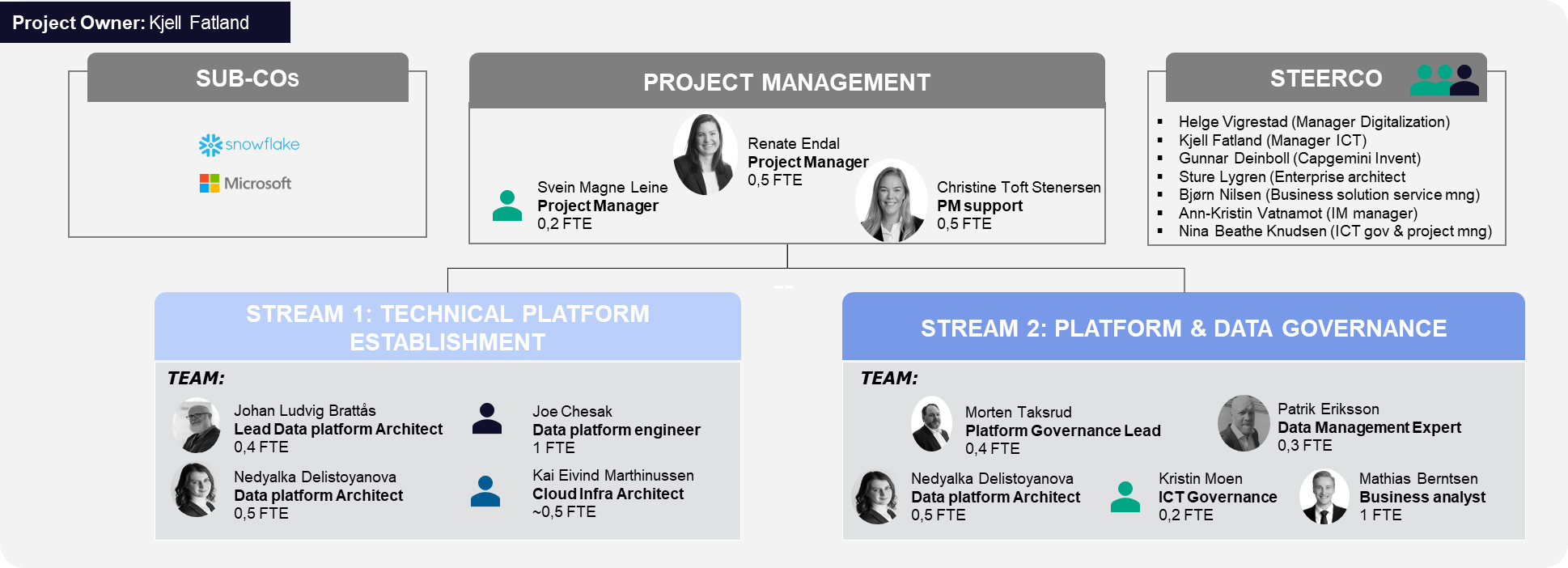
## Governance Structure and Project Processes

**The Data Platform Design Phase Team**

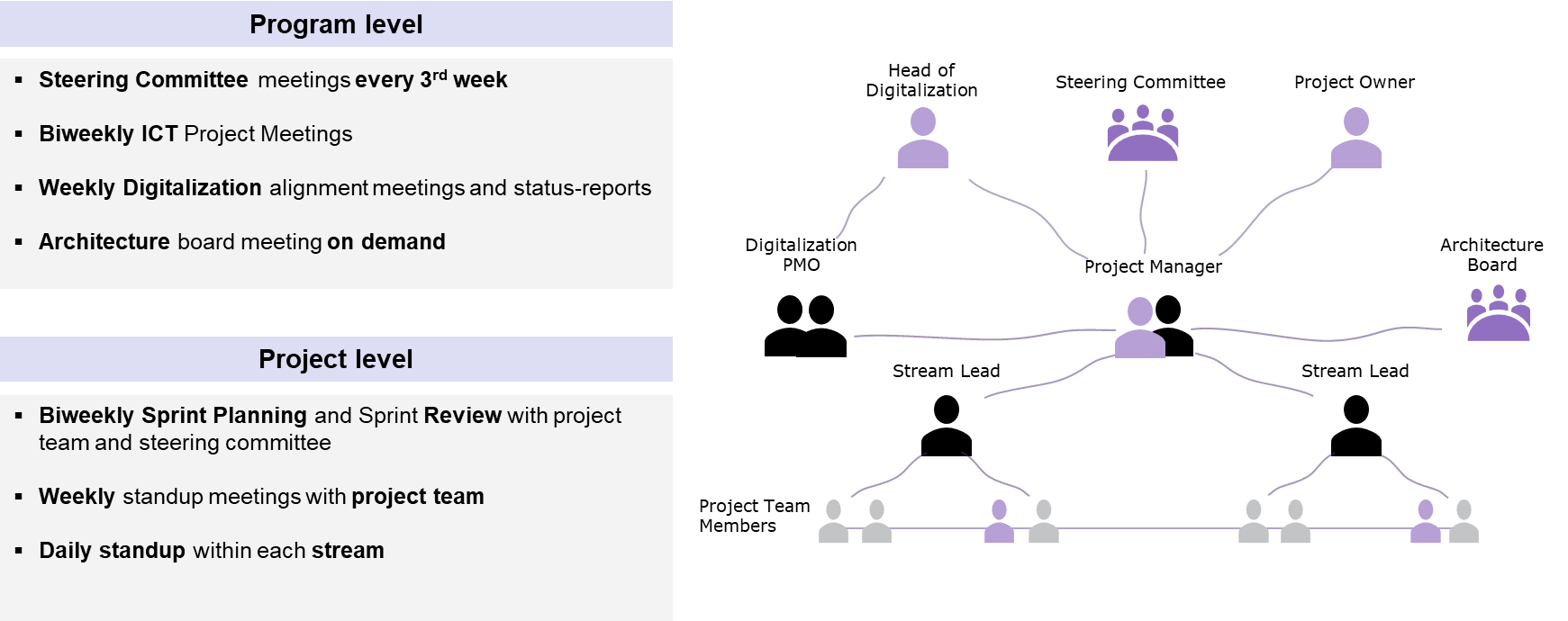


**The Data Platform Implementation Team**

Below is the full Data Platform Implementation Team. It will be “Stream 1: Technical Platform Establishment” who will detail the technical components in the architecture, while “Stream 2: Platform & Data Governance” will focus on Governance Strategy and Processes.



**Communication Overview for data platform implementation, including detailing the architecture**



# Architecture Approach

The project laid out some ground rules of architecture:

* Modular design to make the architecture agile and adaptable.
* As little load on the ICT organization as possible.
* Managed Services where possible.
* Avoid vendor lock-in

## Architecture Process

For the Data Platform project, the Digitalization delivery model is followed: We started out with a Design Phase, where the output was a recommendation of components to include in the Data Platform, in addition to suggestion on vendors for the platform. Further, the Implementation phase aims to detail the architecture based on the recommendation in the Design Phase and Implement an MVP Data Platform. This section contains the approach that was/will be used during these phases.

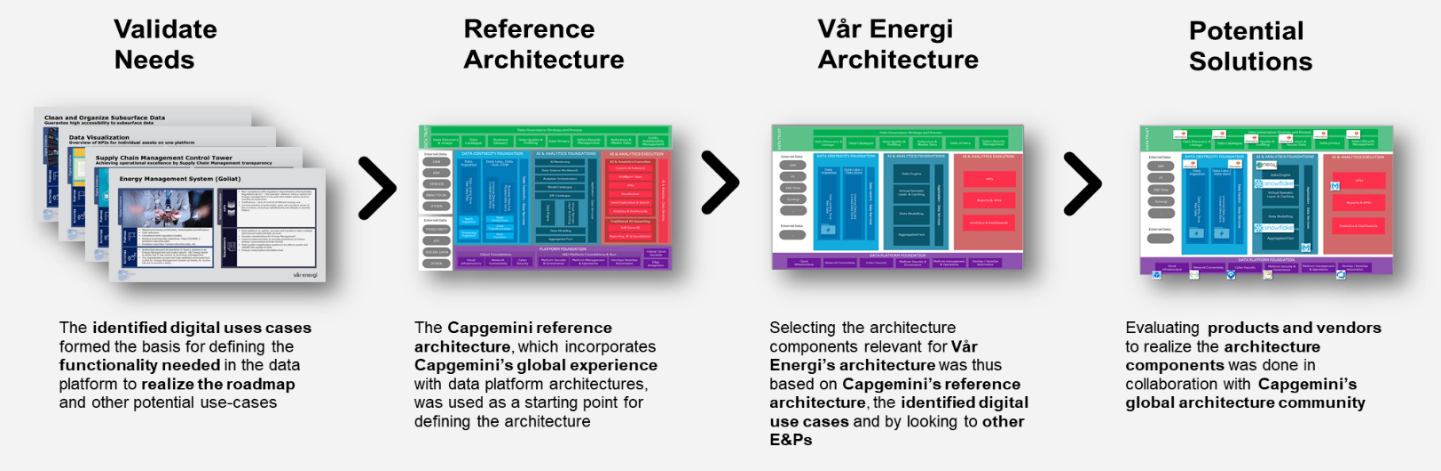
**Approach Design Phase Approach**

As the TOGAF standard was not implemented in the organization, this was not used during this phase.

The point of departure for defining the architecture was the Capgemini Reference Architecture framework. Selecting the architecture was thus based on this architecture, and adapted to the identified needs for Vår Energi and by looking and other E&Ps.

Further, vendors were evaluated in collaboration with the Capgemini Global Architecture community.

The Approach is visualized below

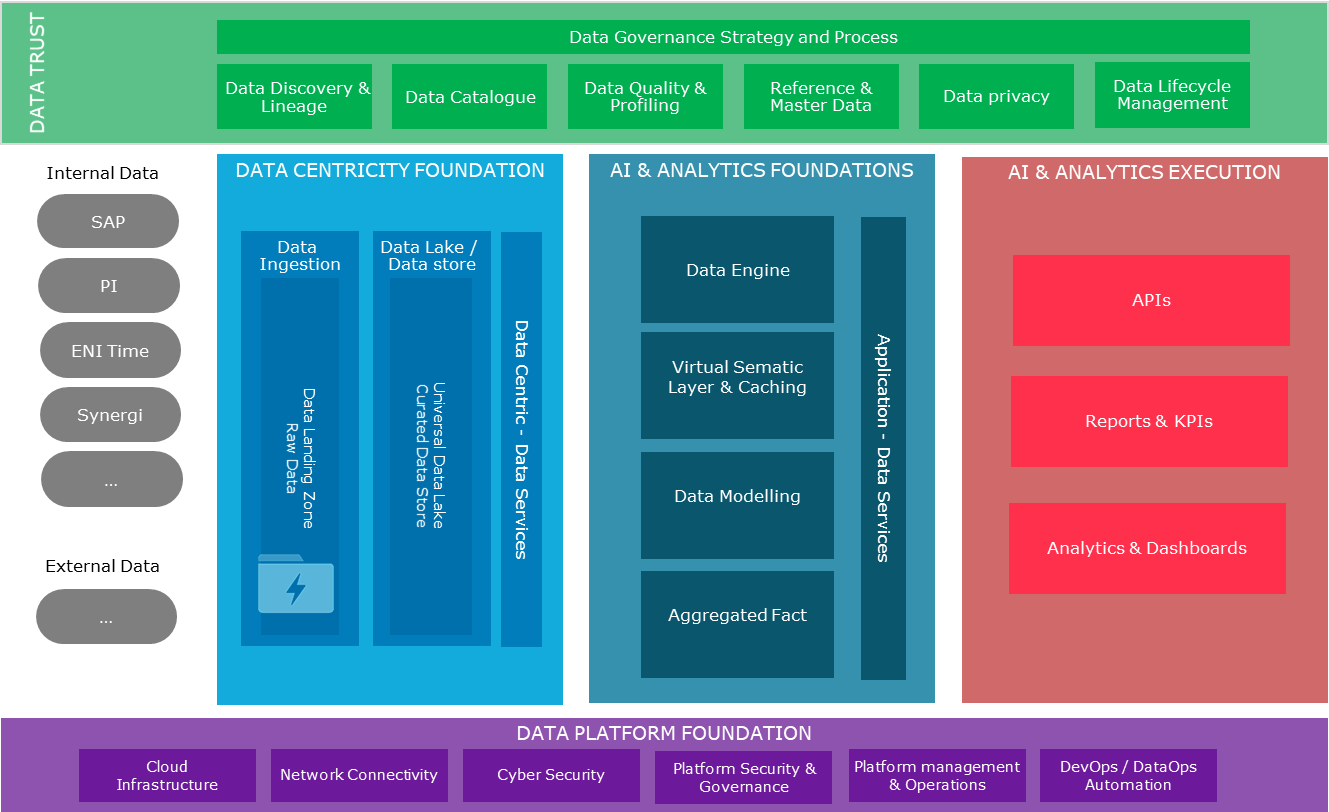


**Approach Implementation Phase / Detailing the Architecture**

The process of detailing the architecture has taken the recommendation from the design phase as a starting point. The scope of the case for how to validate has been determined in dialogue with Vår Energis enterprise architect and has been decided on in the Steering Committee. The detailed architecture has then been defined based on the recommendation and the scope of the case for validation, also taking into consideration the implementation projects timeframe and budget. For Azure, the infrastructure team has been involved in detailing the architecture for the Azure platform. We’ve emphasized that although the platform is primarily set up to support the data platform at the moment, we’ve also taken into account ICTs future plans on how to use the platform to make sure that the architecture choices we make now won’t cause a lot of extra work or put constraints on further development of the platform in the future.

## Architecture Content

***The recommended architecture for the whole data platform at Vår Energi is as follows:***



**Data Trust**

This is where data quality, data governance and is enforced and maintained. As it concerns all stages of the data platform, it spans across all layers.

**Data Centricity Foundation**

While data scientists and certain analysts need access to raw or curated data, most users need access to data that has been prepared and transformed for analytical purposes. This is also the foundation for AI services.

**AI & Analytics Foundations**

Consumption of analytical data and AI processes happens in this layer. Consumers can be dashboards, 3rd party applications, analytical tools.

**AI & Analytics Execution**

A baseline for every data platform is security, orchestration, monitoring, and administration.

**Data Platform Foundation**

The Data Platform Foundation provides services addressing cybersecurity, network, and middleware, furthermore, DataOps, security, governance, and the operating model. Data Platform Foundation ensures that data can be securely ingested from a hybrid cloud architecture, into a secure, industrialized, optimized, automated platform with DevOps tooling.

See section 1.2. in this document for a high level description of the scope of the data platform MVP which will be implemented in the implementation project.

The technical plan document for the data platform goes more into architecture details on the data platform MVP.

## Relevant Methodologies and Industry Standards

Outside of Capgemini’s reference data platform architecture we made use of Gartner’s Magic Quadrants and research reports as well as Forrester’s Wave reports.

## Support of the Enterprise Continuum

As Vår Energi does not have a clear target picture of the overall architecture at this time, it is not possible to say how the Data Platform supports this picture.

However, the needs of the portfolio of digital initiatives have been verified in the process, and thus the Data Platform supports the Digital Portfolio/Strategy.

# Work Plan

The first project set out to identify Vår Energis needs for a data platform and recommend a high-level architecture with recommended products to be use. The first phase of the implementation project will take the recommended architecture as a starting point and agree with the customer upon the scope for the MVP.

## Work Item 01 – Define the Architecture

### Activities

Design phase

* Define scope and technical requirements
* Develop conceptual data landscape architecture
* Design data platform architecture
* Product evaluation and recommended steps

Detailing the architecture in the implementation project

* Meetings with Informatica to get a better understanding of their solution.
* Verify the licensing costs for each of the recommended products
* Based on the input from activity 1 and 2, take it to the Steering Commitee to agree upon scope.
* Get Gartner to give a second opinion on the recommended architecture. Adjust the architecture if Gartners perspectives are very different and ICT wants to do adjustments.
* Detail the architecture based on the agreed scope

### Deliverables

* End-report with architecture recommendation – from the design phase
* Technical plan for the data platform MVP

## Work Item 02 - Technical Implementation

### Activities

We will establish the ingestion from on-premises data sources onto the data platform.

The data centricity layer will be created and the raw data and curated parts of the data lake component of the data foundation. However, no APIs will be stablished for this phase.

The AI and Analytics foundation layer will be established with a single service for querying data, namely the Snowflake data warehouse service.

Data Platform Foundation will be established with the full scope of services needed for the data platform to run as intended, and scale as needed.

The Data Trust layer will have governance strategy and processes established

Activities:

* Creating the data platform infrastructure
* Selecting MVP scope
* Data source discovery
* Creating the data lake
* Creating the Snowflake MVP data warehouse
* Creating the governance strategy and processes

### Deliverables

The following work products will be created as a result of the technical implementation:

* Governance strategy and process documents
* Azure Subscription with networking, security, policies and active directory
* Azure Data Lake Gen 2 (Azure Storage) for raw and curated data
* Azure Data Factory with pipelines for ingesting source data into raw data layer
* Azure Data Factory with pipeline for transforming SAP data into Curated data layer and into Snowflake
* Snowflake data warehouse
* Role-based Access Controls (RBAC) established for all services.

## Work Item 03 – Governance Strategy and Processes

### Activities

Discovery

* Identify and review existing relevant documentation
* Conduct interviews with internal domain experts
* Identify and review relevant external information (e.g. from other companies in the sector)

Development

* Develop the draft for the specified deliverables based on input from the Discovery phase
* Review draft for deliverable with appropriate internal body (e.g. ICT’s Information Management department)
* Review feedback and finalize the deliverable
* Hand over deliverable to the appropriate internal body

### Deliverables

The following work products will be created as a result of the Governance work:

* Naming Conventions
* Data Quality
* Structure and Organization of Data
* Data Lifecycle Management
* Development of Technical Guidelines
* Maintenance and Support Processes
* Request, Approval and Implementation of new reports and data sources
* Define Roles and Responsibilities

## Communications Planning

**Meetings with the Architecture Board**

Initial meeting with the architecture board to go through the recommendations from the design phase.

Meetings with board members to agree on the amount and format of documentation. The documentation will be shared with the board to be reviewed once ready.

After the architecture design has been approved, communications with the architecture board will mainly be on demand or if there are any essential changes to the proposed architecture.

Due to the current situation with many working from home, the meetings will be held digital. If the situations should change during the project duration, we will change the format to in-person meetings.

**Communication and alignment with other stakeholders**

The project will make sure to include relevant stakeholders in concluding and informing about the process of determining the architecture for the data platform MVP. The main stakeholders are outlined in section 3.3. They will be informed by e-mail and we’ll set up meetings to give information an receive input when needed.

## Duration and Effort

**Architecture Design Phase**

**Project Duration**

1. December 2019 - 16. Mars 2020

*13 weeks (excl. Christmas holiday)*

**Effort**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| *Architecture Team* | *Load* | |  | |
| Project Manager | | 20 % | |  |
| Project Member | | 50 % | |  |
| Project Member | | 50 % | |  |

**Detailing the Architecture (Whole Implementation Phase)**

**Project duration**

27.04.2020 - 23.10.2020

CW 18-43 (23 Weeks) \*

*\*Holiday CW 29-31*

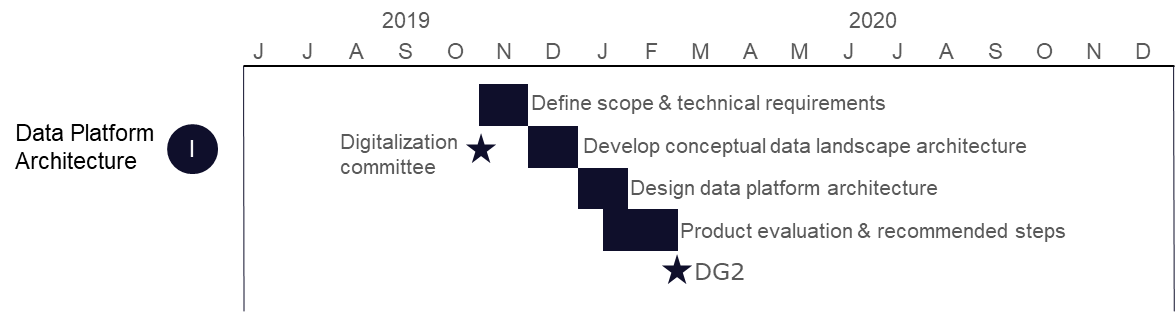
**Effort**

|  |  |  |  |
| --- | --- | --- | --- |
| Role | | Load | |
| Project Management | | 50 % | |
| Project Management Support | | 50 % | |
|  | |  | |
| *Governance Stream* |  | |
| Governance Lead | | 40 % | |
| Data Management Expert | | 40 % | |
| Data Platform Architect | | 50 % | |
| Business Analyst | | 100 % | |
| ICT Governance | | 20 % | |
|  | |  | |
| *Techical Stream* |  | |
| Technical Lead | | 40 % | |
| Data Platform Architect | | 50 % | |
| Data Engineer | | 100 % | |
| Cloud Infra Architect | | 50 % | |

It varies to some extent for how many weeks each of the resources work in the project, but most of them participate throughout the project period.

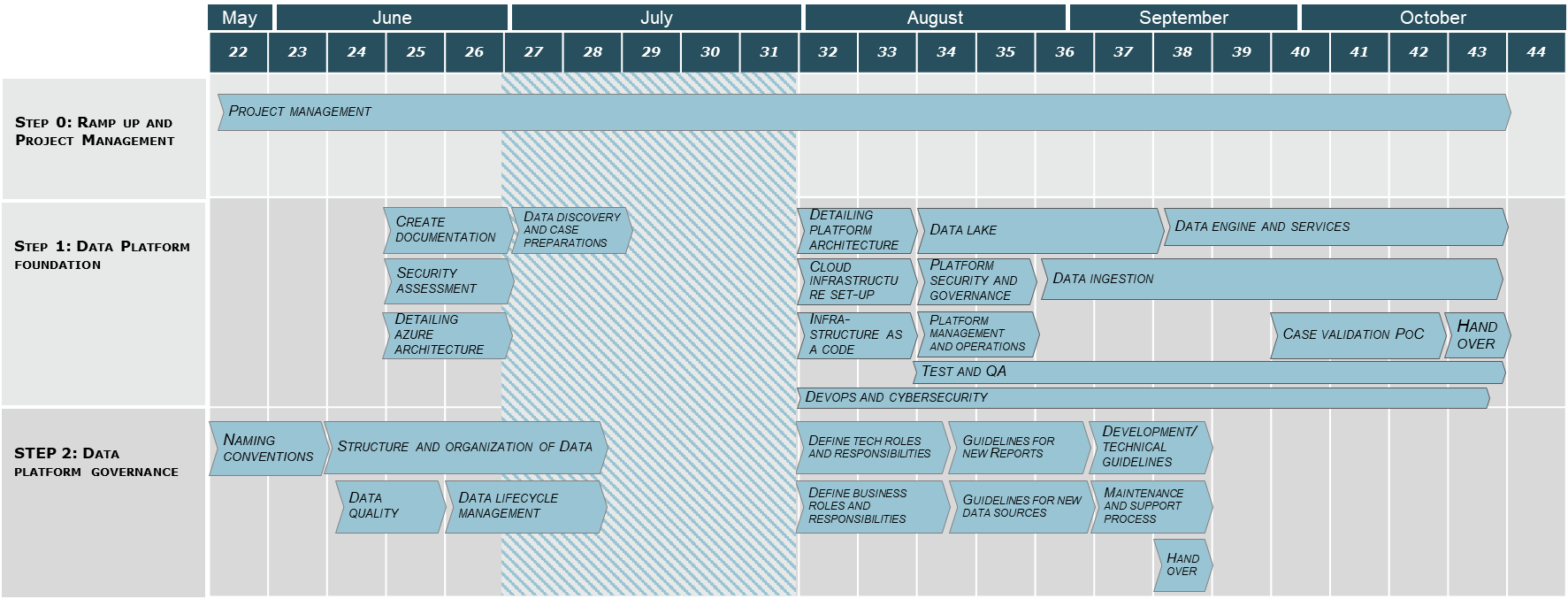
## Project Plan and Schedule

**Project Plan and Schedule for the Architecture / Design Phase**



**Project Plan and Schedule for Detailing the Architecture**

Project plan for the data platform implementation project. The first part of the project will detail the architecture.



# Risks and Mitigations

## Risk Analysis

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **ID** | **Risk** | **Severity** | **Likelihood** | **Mitigation** | **Owner** |
| 1. | Cloud is a new concept for ICT | L | H | Run Azure Fundamentals course for ICT |  |
| 2 | Data Platform stopping up after MVP is delivered due to low readiness | M | L | Make sure to plan and budget project pipeline to make use of established platform |  |
| 3 | Data Trust being neglected | M | M | IM needs to get involved in planning data trust implementation outside of MVP scope |  |
|  |  |  |  |  |  |

## Assumptions

The following table summarizes assumptions for this Statement of Architecture Work:

|  |  |  |  |
| --- | --- | --- | --- |
| **ID** | **Assumption** | **Impact** | **Owner** |
| 1. | This statement of architecture work is filled out based on the information currently available. The statement may be subject to change at a later stage as we proceed with the implementation. | The statement may need to be changed | Renate Endal |

# Acceptance Criteria and Procedures

## Metrics and KPIs

The Implementation phase of the Data Platform will deliver a MVP / PoC of the architecture, and is thus a validation of the architecture. As such, the deliverable of the Implementation project will implicitly show if the architecture is a success or not.

## Acceptance Procedure

This will be provided when the procedure is set by the Architecture board.

# Signature Approvals

|  |  |
| --- | --- |
|  |  |
| Signature | Date |