First Draft

The Danish Case:

Improving water resilience and competitiveness through an integrated, cost-effective approach to data and IT

Figure 1: European Water Data Ecosystem

Table of contents

Background ... 3

Strategy for DMP ... 5

- Summary ... 5
- Business drivers ... 5
- Principles ... 8

Business architecture ... 10

- Summary ... 10
- Stakeholders ... 10
- Capabilities ... 11
- Business processes ... 13

Data architecture ... 16

- Conceptual data model ... 17
- Data models ... 18
- Entity definitions ... 23

Technological architecture ... 25

- Summary ... 25
- Solutions ... 27
- Components ... 29
- Data sources ... 30
- Data quality ... 30
- Target architecture (DRAFT) ... 31

Background

This document is intended to outline how Denmark is addressing water resilience competitiveness through an integrated, cost-effective approach to data and IT.

In 2007, the Danish Environmental Portal (DMP) was established as a central organisation with the responsibility to maintain and operate the environmental data on behalf of the Danish state and its municipalities.

2007 through 2009, DMP assumed ownership of several legacy solutions used for research and monitoring of water data. Until then, the development and operation of the legacy solutions were owned by municipalities and the Danish state to solve specific and individually purposes.

Throughout, the establishment of the Danish Environment Portal, the key focus has been simplification of the water environment ecosystem to build a competitive edge for the users of the water and chemistry data in Denmark.

A cornerstone of the water environment ecosystem is a national portal for water and chemistry data ranging from groundwater, surface water, marine water, drinking water bathing water to urban wastewater available for everyone. The national portal is supporting research, monitoring and reporting, and decision making.

The efforts have proven successful in many ways. However, new needs to be addressed continuously. At the same time, technologies are evolving with new and better ways to enable the needs.

The key messages from this document are summarised below:

- Focus on data use
- Focus on speed
- Interoperability at source
- Simplify data models
- Access to raw data

Data collection prioritized by the use cases optimises real value creation, as getting the right data is more important than getting all data.

Competitiveness is when decisions are based on accurate and relevant data, hence it is essential that time from data collection to data usage is as short as possible.

Interoperability at source increases competitiveness as data sharing of raw data enables a broader usage, shifting from single purpose to multi-purpose solutions. Access to the raw data is necessary, since transformed data loses critical information.

As results, the following insights summarises key recommendations:

Secure funds to maintain a competitive edge through exploration of modern technologies e.g., usage of AI. Some of the challenges that DMP is facing as a semi-public organisation, are funding continuous development and operation of the data ecosystem. Funds are usually granted to archive political initiatives and target specific outcomes, but the data ecosystem serves a wider purpose.

Implement a well-defined data governance model for data sharing across different solutions across public offices and entities. It requires effective processes and clearly defined responsibilities for data reporters and data owners to ensure accuracy and integrity of the data continually e.g. data should only be soft deleted to keep the historically reference and all changes to the data should be logged.

Establish an enterprise architecture function with the mandate to ensure standardisation and simplification within the organisation. Further, the enterprise architecture function should be promoting components to be shared more widely.

Design a common data model for water and chemistry data to reduce complexity. Shared environmental data with interoperable water and chemistry data enables generic processes resulting in a more cost-effective water ecosystem. The Jupiter, VanDa and Puls solutions share 80% of a common data model.

Develop standard components to accelerated harmonisation of water and chemistry data through interoperability. Parameterlist has resulted in cost savings in efficiency in administrative tasks related to chemical measurements 28-84 million DKK[1].

[1] Report - Value estimation of a Standardized Chemical Parameter System - version 1.0 - 06-11-25

Strategy for DMP

Summary

DMP is working with a set of strategic arenas where initiatives are expected to create value across the organisation. The goal is not only delivering projects on time and within budget, but generating synergies across domains.

There is a shared recognition that no “perfect” organisational model exists; instead, the aim is to strike an effective balance between product-oriented delivery and strategic anchoring with stakeholders and funding partners.

The strategic focus is on establishing a common direction and coherence across the value chain, including alignment with EU directives and environmental objectives.

There is a clear strategic shift away from siloed projects towards shared architecture and platform thinking across systems e.g., Vanda, Puls, Jupiter.

There is strong emphasis on:

- standardisation over local optimisation
- reuse of components
- establishing a shared roadmap towards the target architecture

External collaboration is conducted through a small number of key ambassadors who possess both domain expertise and digital understanding.

The strategy is shifting from project-driven delivery towards platform- and data-driven governance, with a strong focus on shared standards and cross-organisational value creation.

Business drivers

Competitiveness

Enable high-quality, data-driven decision making across public and private sectors by ensuring authoritative, transparent, and accessible water and chemistry environmental data.

Goals

- Publicly funded data is consistently registered in authoritative solutions across relevant organisations such as universities, DTU, Fehmarnbelt, municipalities, regions, other public and private organisations.
- Data refinement occurs on a shared, transparent ecosystem where authorities retain ownership.
- Data and index calculations are relevant, accurate, accessible, and easy to understand.

- Decision making processes for public and private organisations needs are supported by complete, consolidated information.

Outcomes

- Higher quality decisions due to consolidated, authoritative data.
- Immediate public accessibility of publicly funded data.
- Increased usage driven by improved accessibility and clarity.
- Continuous monitoring capability for environmental measures, including aquatic environment improvements.

Simplification

Reduce complexity and cost by consolidating investigations, workflows, and user interactions into a shared water environment ecosystem.

Goals

- New investigations e.g., HSD, ODA-Exit are developed within the ecosystem.
- Technology encourages direct data registration through improved user experience, including tablet and app-based solutions.

Outcomes

- Lower development and maintenance costs due to fewer solutions.
- Functionality is developed once and reused across the ecosystem.
- Public and private organisations avoid developing and maintaining their own solutions.

Standardisation

Create a coherent, interoperable ecosystem for authoritative water and chemistry environmental data through consistent processes, systems, and data structures.

Goals

- Establish a unified ecosystem for all authoritative environmental data.
- Ensure consistent use of authoritative solutions across stakeholders.
- Handle identical investigations e.g., chemical analyses in a uniform manner.
- Improve coherence between current solutions: VanDa, PULS, Jupiter, Results Database, Water Plan Data, Environmental Data, Land Use Data.

Outcomes

- Reduced training needs due to standardised processes and interfaces.
- Faster task execution through consistent workflows.
- Portfolio-wide process improvements and shared benefits.
- Higher data value through consistent collection and easier data integration.
- Improved data quality through uniform storage and governance.

Principles

These architectural principles should be followed to improve water resilience and competitiveness through an integrated, cost-effective approach to data and IT.

	Statement	Rationale	Implications
Efficiency	Monitoring and software development must be executed in a cost-effective, resource-efficient manner using simplified, standardised processes and common components	Common data model and components reduce duplication, lower costs, and streamline operations	Processes must be standardised across organisations Development must prioritise reuse Shared ecosystem becomes the default IT environment
Accessibility	All collected data, including underlying datasets and index calculations, must be freely accessible to all relevant organisations and persons	Open access maximises the value of publicly funded data and supports transparency, collaboration, and informed decision making	Data must be published without unnecessary restrictions. Systems must support easy discovery and use of data. Access policies must be consistent across the ecosystem.
Relevance	Data must be accurate and fully transparent from collection to final use.	Complete, usable and trustworthy data is essential for environmental management and decision making.	Quality assurance must be embedded consistently throughout the value chain. Metadata must document provenance, methods, and assumptions. Index calculations must be traceable and reproducible
Interoperability	Data must be integrable and aligned with	Standardised data enables cross-system	Solutions must adopt shared the common

	Statement	Rationale	Implications
	national and EU standards, including common code lists	integration, reduces complexity, and supports national and international reporting requirements	data model and code lists. Interfaces must support standard exchange formats. Legacy solutions must be adapted or decommissioned to ensure compatibility
Open development	All code, algorithms, and index calculations must remain under public ownership and be developed in the shared ecosystem	Open development increases transparency, reduces vendor lock-in, and enables reuse across the ecosystem	Development must be on shared repositories. Outputs must be published under open-source-compatible terms and licenses. Authorities must have stewardship of all developed assets
Data governance	Authorities[2] are responsible for refining water and chemistry environmental data, while DMP manages collection, storage and presentation of both raw and refined datasets	Clear governance roles ensure accountability, data quality, and consistent stewardship across the value chain	Governance responsibilities must be formally assigned and maintained Refinement processes must follow agreed standards and controls DMP must provide authoritative access to all datasets

Table 1: Architectural principles

[2] Ministry of Environment, Environmental Protection Agency, Agency for Green Land Conversion and other authorities

Business architecture

Summary

The business perceives itself as highly differentiated; however, analyses show that processes and needs are approximately 80% identical across the organisation.

There is a need for:

- shared business models
- harmonised use cases and user stories
- a clear understanding of intents and needs before solutions are designed

A central challenge is that development activities often proceed without a shared and coherent understanding of:

- the relevant use cases
- the expected value creation

which results in costly and fragmented development efforts.

The organisation is working towards:

- shared components e.g., rules engines, planning modules, permissions
- cross-domain reuse across water, chemistry, and environmental domains

From a business perspective, new requirements are driven by:

- increased monitoring obligations such as EU directives
- the need for continuous insight into environmental data
- integration of a broader range of data sources e.g., sensors, satellites, laboratories

The business architecture is moving towards shared, cross-cutting capability models and data-driven decision support, yet progress is still hindered by persistent silo thinking.

Stakeholders

With the establishment of DMP, the governance model changed from a few local owners to a central ownership that represents the regions with 10%, the state with 45% and the municipalities with 45%. The state is represented by Danish Ministry of Environment and the Danish Environmental Protection Agency.

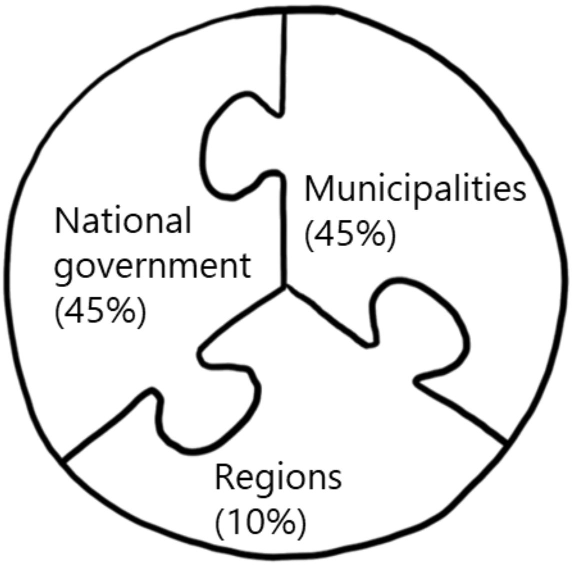

Figure 2: Shared ownership

The governance is supported by the advisory groups listed below.

- Land
- Nature
- Surface water
- Urban wastewater
- Groundwater
- Rat control
- Environmental impact assessments
- Data sharing

The data platform has many different players ranging from data reporters, owners, and end-users. The data reporters and owners are various businesses, research institutions, and public offices[3]. The advisory groups are acting as representation for end-users.

Capabilities

Each of the legacy solutions were developed to collect and manage water data with for a specific purpose as listed below:

- Data collected as field samples and measurements in lakes, streams, and marine areas and the results from the analyses at laboratories.
- Data on the animal life in lakes and marine areas.
- Data on environmental contaminants in lakes and marine areas.

[3] De Nationale Geologiske Undersøgelser for Danmark og Grønland (GEUS), Danish Centre for Environment and Energy (DCE)

- Data about fish in lakes and marine areas.
- Data about vegetation in lakes and marine.

Water data was originally collected and used to research purposes only. Later, the water data was used to monitoring environmental changes in lakes, streams, and maritime areas and to reporting to the EU. Today, the water data is increasingly used to making data driven decisions related to agriculture and climate initiatives. In near future, the water data might be used to forecasting the potential effects of these initiatives.

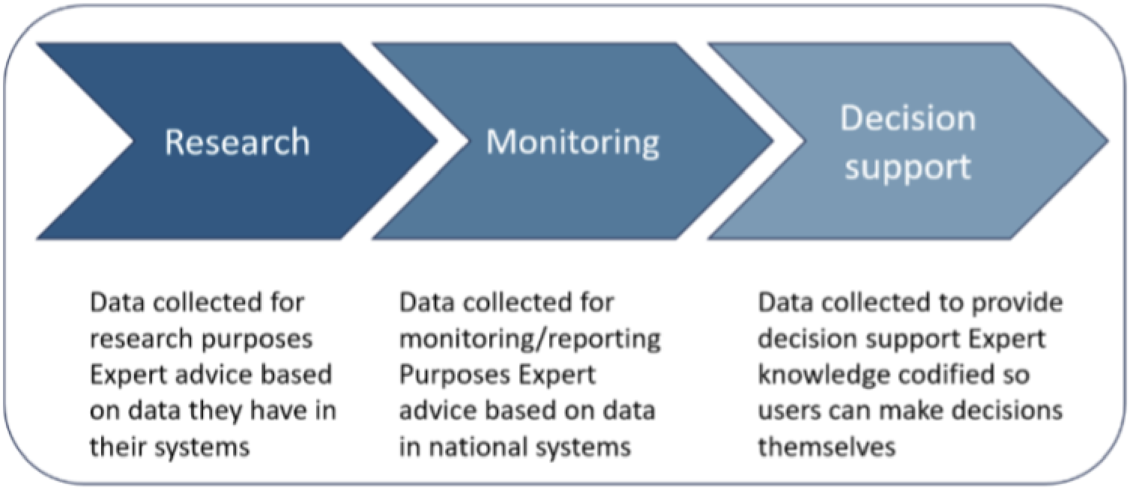

Figure 3: Process model

Water data is defined as groundwater, surface water, marine water, drinking water bathing water and urban wastewater[4].

- Groundwater is fresh water stored underground in aquifers.
- Surface water is fresh water located on top of land, forming terrestrial waterbodies such as rivers, lakes, reservoirs, and wetlands.
- Marine water is surface saltwater within one nautical mile from coastal line and includes lagoons and salt marshes.
- Drinking water is water safe for ingestion.
- Bathing water is water suitable for bathing through monitoring for safety and quality.
- Urban wastewater is water from households and industries, containing pollutants.

DMP identifies surface water and marine water collectively as surface water.

Figure 4: Water cycle

Business processes

The water data is used to reporting and monitoring the conditions and development of environmental situation within lakes, streams, and marine areas. The water data is reported to various EU agencies. The water data is used for the administration and control of permits of discharge of water; stream restoration and regulation, and to develop water and marine strategies in Denmark.

While the reporting, monitoring and querying processes were standardised across the users and aligned with the new data model, the new data platform was required to remain compatible with the current data sharing processes.

[4] Water - Environment - European Commission

One of these processes provides the database ObservationData (ODa) with water data which is used for research purposes by Aarhus University and Danish Centre of Environment (DCE).

Another process stores data related to areas in static views enabling end users to query and visualise the area data.

Further, the new data platform integrates with the current StanLab solution used by laboratories to record samples and their results.

The solutions EA-Hub and EA-Tools are using surface water data from the new data platform to generate environmental impact assessments.

The benefits have increased as the data has been standardised and harmonised with a national water portal and a common data model. Some samples of how the water environmental data has benefitted national wide purposes beyond the original needs for monitoring and reporting are listed below.

- Danish Meteorological Agency (DMI) uses the national water data to flood warnings.
- Banedanmark collects data for water levels near their rail tracks to prevent floods damaging the tracks.

The table below displays key needs addressed by Danish authorities using the water environmental data.

<table>
<tr><th></th><th>Planning</th><th>Execution</th><th>Follow-up</th></tr>
<tr><td>State</td><td>Establish landfill control program

Establish groundwater monitoring program</td><td rowspan="3">Make measurements and observations in the field

Take samples

Perform laboratory analyses</td><td rowspan="3">Receive warnings

Identify exceedances

Analyse results

Issue new injunctions

Decide next actions</td></tr>
<tr><td>Regions</td><td>Implement projects at pollution sites

Establish control plans for abatement facilities</td></tr>
<tr><td>Municipalities</td><td>Issue control programs for water supplies

Establish inspection plans for energy facilities</td></tr>
</table>

Table 2: Key needs solved with water environmental data

The shift to an interoperable data model with shared components has resulted in increased value through simplification and more generic processes.

Typical processes that have improved by the implementation of a modern water ecosystem are the following.

Processes towards EU:

1. EU writes a directive
2. The Danish Environmental Protection Agency (DEP) transposes the directive to Danich legislation and regulation
3. The regions and municipalities receive the Danish legislation and regulation
4. The DEP retrieves data from the Danish authorities
5. The DEP must report data to the EU

Processes in Denmark:

1. Public authorities must establish monitoring and control programmes

2. Waterworks and wastewater facilities must investigate the substances as specified in the programmes
3. The laboratory delivers results to the public authority and reports them to the mandatory reporting system
4. The responsible authority must perform quality assurance of the data
5. The authority must process data according to applicable regulations such as assessing measurement results against current threshold values

Data architecture

Summary

Significant challenges include:

- lack of standardisation in code lists and data structures
- multiple parallel “truths” e.g., three different rules engines

There is a need for:

- a national data foundation
- shared data models and common semantics
- open and accessible datasets

Current issues:

- data is partially “locked” within systems e.g., Vanda
- insufficient governance of standards
- weak alignment between data capture and data publication

Ongoing initiatives include:

- data marts for data exposure though combining datasets remains difficult
- event-based data handling
- shared parameters and code bases

Future requirements involve:

- managing large data volumes e.g., non-target analyses, sensor data
- combining:
    - quality measurements
    - quantitative volumes

Key focus areas:

- making data more open and accessible including for AI use
- ensuring consistent data across actors and countries

The data architecture is transitioning from a fragmented landscape to a standardised, shared, and scalable data platform, but still lacks sufficient governance and harmonisation.

Conceptual data model

Each of the legacy solutions were developed with different data models and naming conventions. For the data to become interoperable, common standards and a common data model were required.

A common data model to make water and chemistry data shareable was developed using the observations and measurements reference architecture[5] based on the "Observations, measurements and samples" concept, which is originally described as ISO-standard (ISO 19156) and updated in 2023.

The concept is a generic model used when it is needed to know something about someplace. The concept does not only apply to water properties but can be used for all kinds of data. E.g., if data about the waterflow of a river is needed, waterflow is defined as shareable property, and how it is measured. Once defined, the description of the property waterflow can be applied for all rivers, enabling data interoperability.

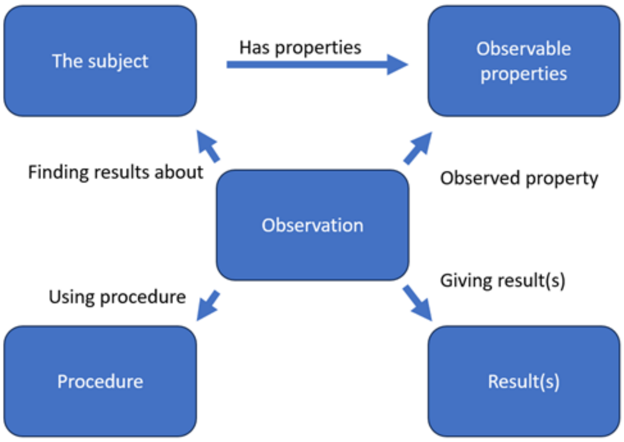

Figure 5: Observations, measurements and samples concept model

Observable properties

All physical objects have properties that we can measure. E.g., all rivers have the same properties such as waterflow, temperature, water level etc. Each observable property is

[5] Referencearkitektur for observation og måling | Fællesoffentlig Digital Arkitektur

given a semantic definition, a unit of measure, an observation procedure on how to measure and a constraint for the result.

Observation

A measurement taken at a specific time that falls within the permitted range for the observed property, following the agreed procedure and assigned data quality.

Procedure

Each observable property has defined procedures for how it can be measured. When an observation is performed, the specific procedure used is identified, even when multiple methods are available.

Result

The result is the product of the observation.

Data models

As common data model was conceptual and generic, it was too abstract from the current data models used by the researchers. The common data model was revised to address the gap and enable the development of common standards.

During the interviews of the researchers and users to understand the collected data and its usage, it was discovered that some data was never used. Hence, the common data model helped to reduce the complexity by only collecting data that was used.

Furthermore, the data was standardised and harmonised by sharing code lists across public offices.

The logical data models implemented with the solutions Puls and VanDa are based the common data model.

The data model for the solution Puls includes the water and chemistry related data domains:

- Wastewater treatment plants for municipal and industrial sewage treatment facilities
- Outfall for combined sewer overflows, stormwater discharge points and retention/detention basins
- Bathing water stations for the EU Bathing Water Directive 2006/7/EC compliance monitoring stations for public beaches

The data model for the solution VanDa includes the following data domains related to water and chemistry:

- Type Lists
- Water Area
- Field Measurements
- Field Observations
- Measuring Depths
- Sediment
- Water Chemistry

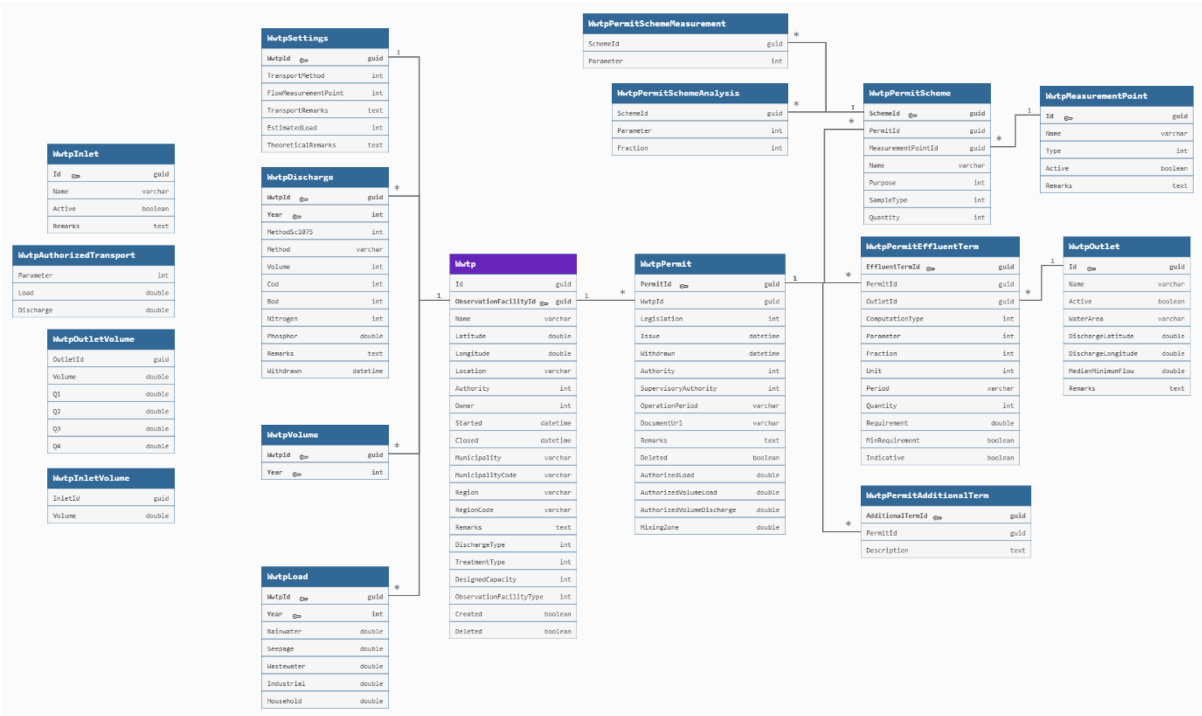

Figure 6: Wastewater treatment plants data model

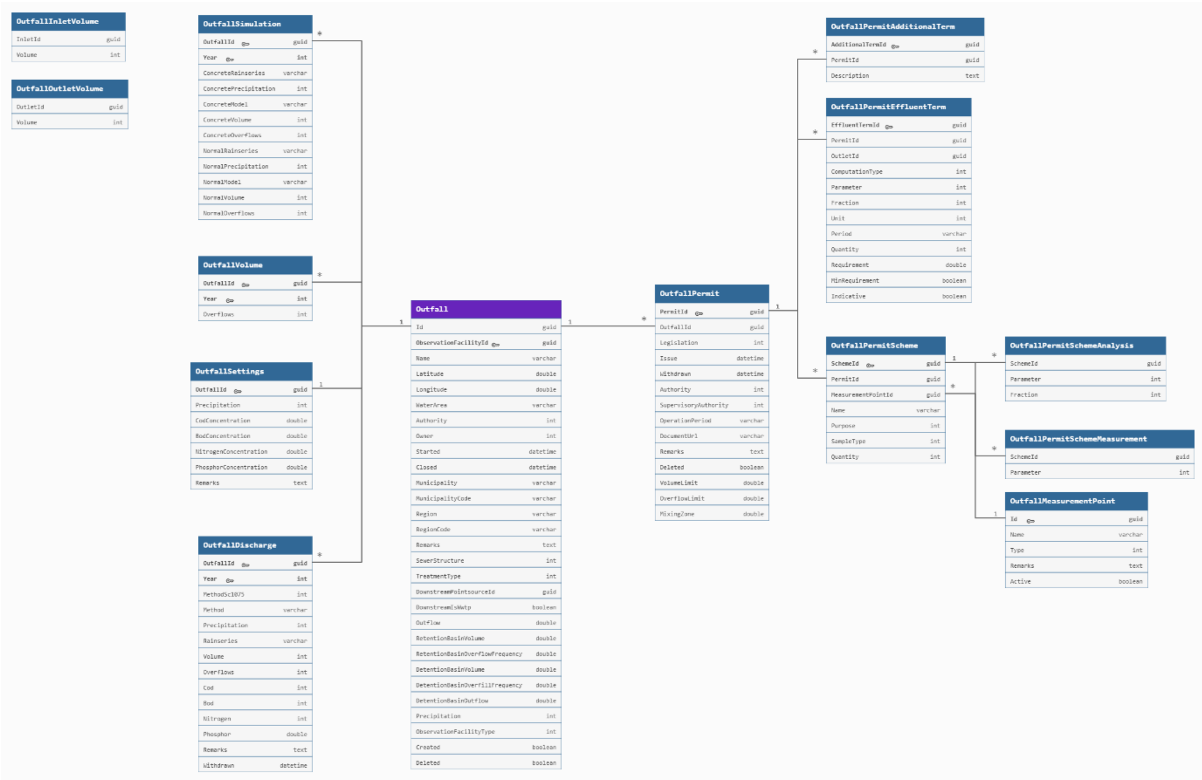

Figure 7: Outfall data model

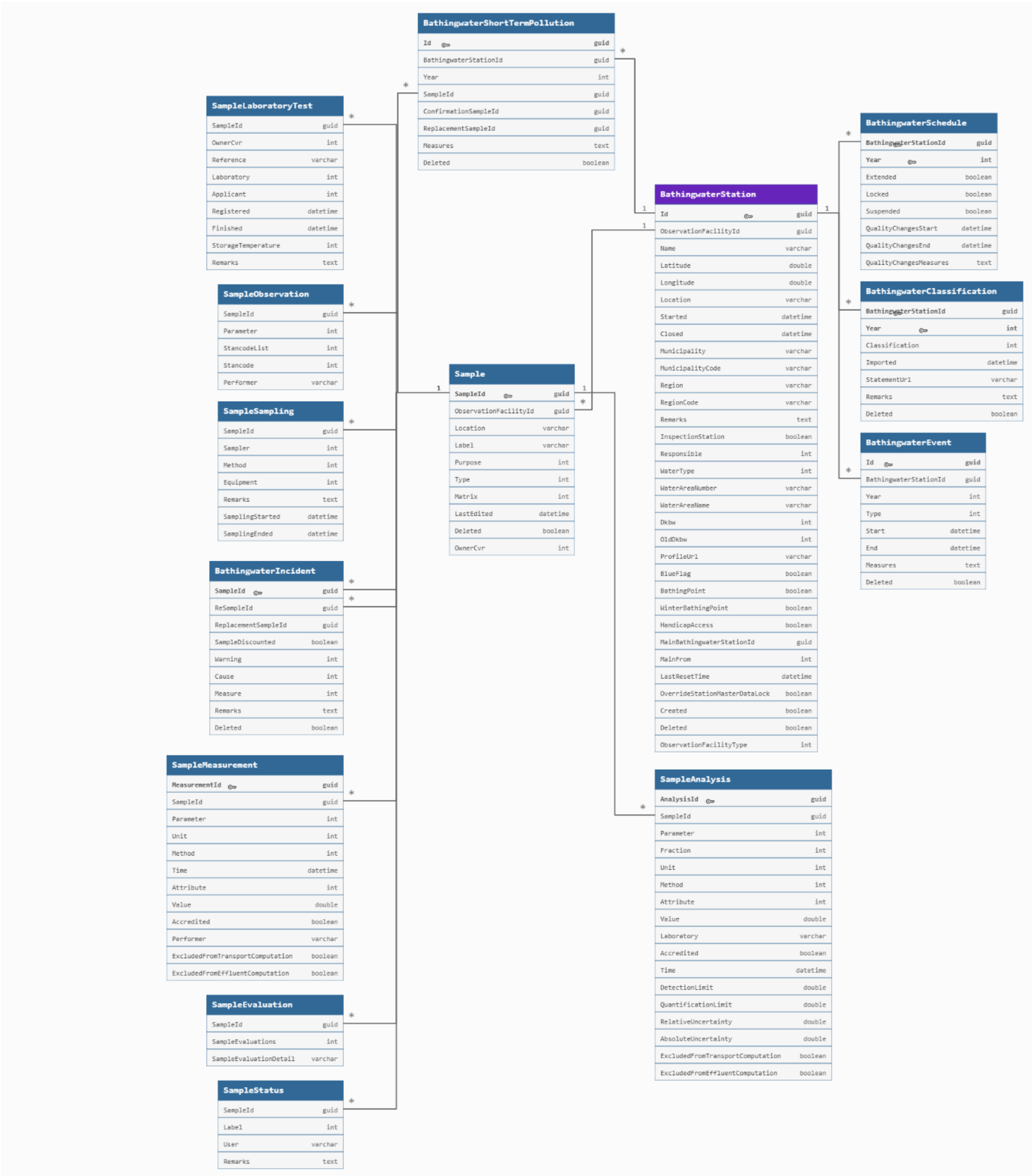

Figure 8: Bathing water station data model

Figure 9: Water chemistry data model

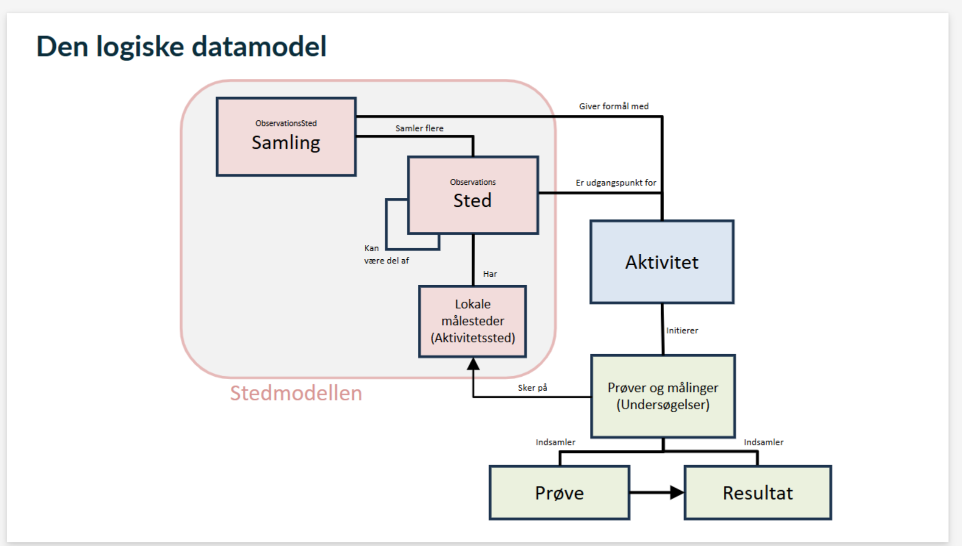

Figure 10: Jupiter data model changes

Entity definitions

Entity	Entity definition
Collection	Grouping of one or more Facilities that together represents a real-world entity that is treated as a whole. E.g. a water treatment plant is a collection that comprises of many stations which are all the individual locations in the water treatment plant where samples may be taken
Station	Information about where the activity has taken place
Activity	Information about the activity that has taken place on a specific location, and for a specific purpose
VanDa Type List	Local code list or a shared code list
VanDa Water Area	Location of the test of soil water using a suction lysimeter
VanDa Water Area	Location of the test e.g. freshwater lakes, streams, rivers and the sea
VanDa Subject Catalogue	Relationship between substances, units and examination subjects
VanDa Field Measurement	Measurement in the field rather than in the laboratory e.g. temperature and pH values
VanDa Field Observation	Visual observation about the environment near streams e.g. trees, and rocks, or similar, recorded in the field in connection with the sample
VanDa Measuring Depth	Information about e.g. “view to bottom” and thermocline in lakes and the sea
VanDa Sediment	Chemical analyses of sediments from the bottom of streams, lakes and the sea
VanDa Water Chemistry Examination	Information about the presence of environmentally hazardous substances in water environments. Samples

Entity	Entity definition
	are collected for subsequent chemical analysis, after which the samples are submitted to laboratories that provide the final examination results
VanDa Water Chemistry Sampling	Information about how the test was taken e.g. equipment, and the start and end time
VanDa Water Chemistry Sample	Information about how the test was taken e.g. by type, and water depth
VanDa Water Chemistry Subsample	A smaller, clearly defined portion of a sample
VanDa Water Chemistry Result	Chemical test result from the laboratory
Puls Bathingwater Station	Main entity representing bathing water quality monitoring station
Puls Bathingwater Schedule	Annual sampling schedule for bathing water quality monitoring
Puls Bathingwater Classification	Annual water quality classification per EU Bathing Water Directive
Puls Bathingwater Event	Event affecting bathing water station operations and public access
Puls Bathingwater Short Term Pollution	Predictable short-term pollution event managed separately from incident
Puls Bathingwater Incident	Pollution incident requiring management actions and public notification
Puls Sample	Information about the test of water quality taken at the bathing water station such as the collection of observations and analyses on a specific date
Puls Analysis	Analysis describes the different substances analysed and the results from the laboratory
Puls Sample Laboratory Test	Laboratory test contains information about the laboratory analysing the sample e.g. who performed it, when, etc.
Puls Observation	Qualitative assessment and visual observation, e.g. number of bathers, or similar, recorded in the field in connection with the sample
Puls Sampling	Information about how the water quality test was taken e.g. by whom, method, and the start and end time
Puls Measurement	Measurement in the field rather than in the laboratory
Puls Evaluation	Quality control evaluation stamps for entire samples
Puls Sample Status	Whether the sample is planned, collected, in the laboratory, completed or cancelled

Table 3: Selected definitions of entities in Puls and Vanda

Technological architecture

Summary

Strategic technology choices:

- cloud-native PaaS as a foundational principle
- centrally governed architecture instead of outsourced to vendors

Architecture patterns:

- event-driven architecture
- separation of:
  - data capture
  - data exposure (read/write separation)

Key focus areas:

- lowering complexity by reducing the number of components
- platform-level standardisation
- governance of technology choices

Current challenges:

- historical technical debt
- inconsistent solutions across projects
- lack of a unified UX/UI design approach
- insufficient testing and quality assurance

Operations and development:

- DevOps-inspired delivery teams
- need for strong platform governance
- cloud optimisation has significant cost impact

Forward-looking priorities:

- shared components
- consistent user experience
- improved automation and testing
- high scalability and performance

The technical architecture is modern with cloud-based and event-driven, but is hindered by inconsistency and insufficient standardisation across solutions.

The legacy solutions were developed for specific purposes by different vendors with different technologies. The operational costs and conditions of the legacy solutions led DMP to replace them with a modern cost-effective data platform.

The new data platform was designed according to the cloud-native principle using standard cloud computing and cloud storage services to ensure scalability, reliability, and security. The new data platform was required to be able to create 1000 transactions per second and manage historical data in two timelines known as bitemporal modelling.

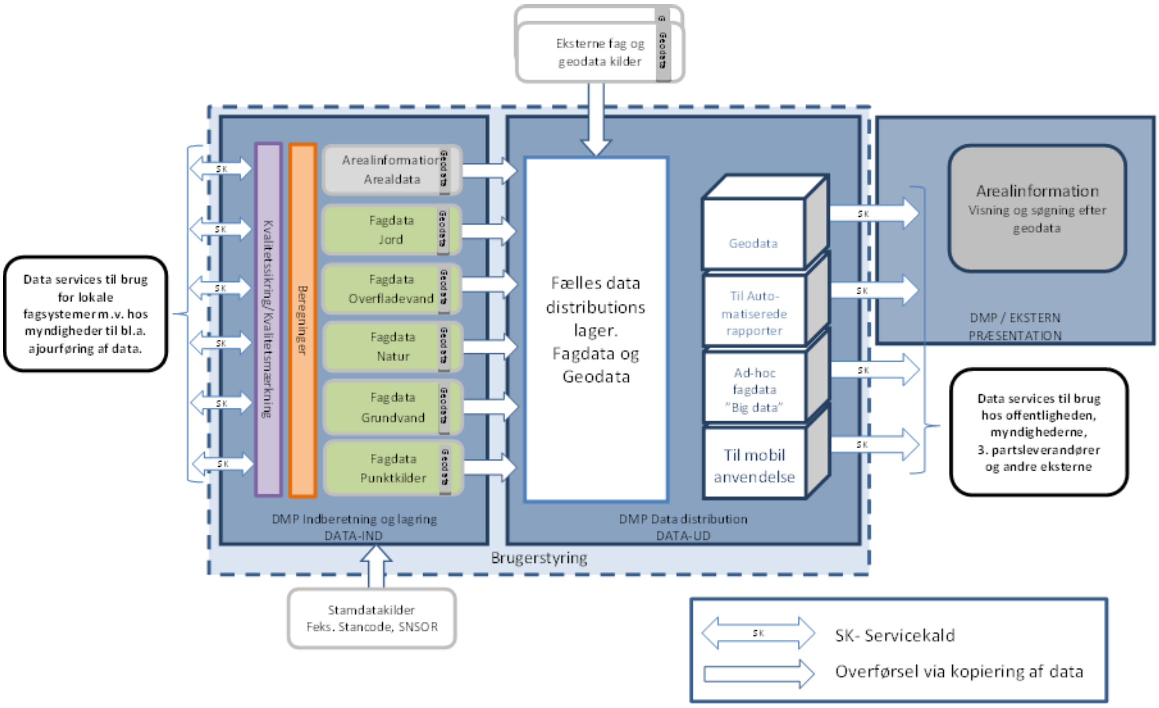

Figure 11: Technological architecture

The new data platform was required to adhere to three key principles in the target IT architecture:

1. Separate data input and data output
2. Use a standard setup for cloud computing
3. Leverage standards and shared components

Further, the data platform was given the following requirements through the target IT architecture:

- It must be a robust, scalable, and secure environment
- It must provide access to data for applications through via loose coupling such as RESTful API
- It must provide easy and integrated access to all the data within the Environmental Portal

- It must ensure standardised and harmonised data across domains
- It must enable filtering of data by quality classification
- It must enable connectivity to data to standard OTS software such as Microsoft Office
- It must provide access to extraction of large data sets, usage on mobile devices and AI/ML execution
- It must enable connectivity to internal and external through a secured gateway with the option for payment

To benefit from the target IT architecture, the data needs to be standardised and harmonised. The distribution data needs to be cleansed for any personal information (GDPR), hence the need to have data governance in place.

The transition to the target IT architecture was done incrementally as new business needs were presented.

The development process to enable the new business needs follows the 5-step approach:

1. Identify and prioritise the new requirements
2. Analyse and select the options using standard component and in-house development guided by the target IT architecture
3. Develop a Proof of Concept to pilot the implementation
4. Evaluate and document the pilot for realised benefits
5. Scale the pilot

Solutions

Since 2016 DMP has been developing and operating a data platform for all environmental data in Denmark. The purpose was to replace the legacy solutions with a single modern cost-efficient data platform. The new data platform was based on standardisation and harmonisation of data to ensure interoperability across the various usages.

Today, the water ecosystem is consisting of multiple solutions with shared components. This is a result of DMP has continually migrated 36 systems into a modern water ecosystem for water and chemistry data with the following key solutions:

- Jupiter, a database of 418.000 sampling locations in Denmark
- PULS, a solution for monitoring and reporting of water discharge from sources like wastewater treatment plants, rain-induced outlets, aquaculture facilities, and bathing water

- VanDa, a solution for collecting and managing data about the Danish water environment

The solutions VanDa and Puls are developed with an event-sourced Command Query Responsibility Segregation (CQRS) architecture with separation of databases for writing and reading. Further, VanDa and Puls are implemented with infrastructure as code open-source software to standardise the cloud-based technology stack across the development projects at DMP.

The following diagrams display the system architectures.

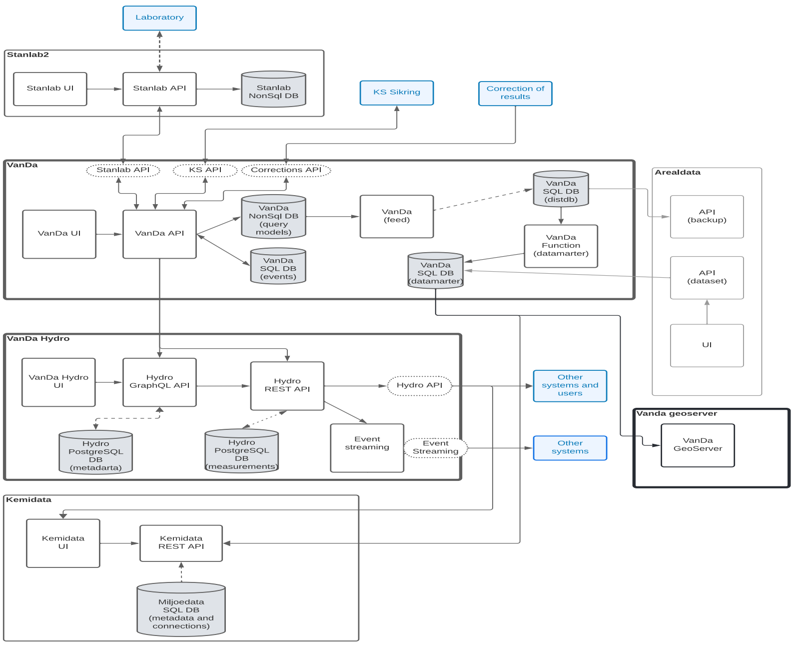

Figure 12: VanDa system diagram

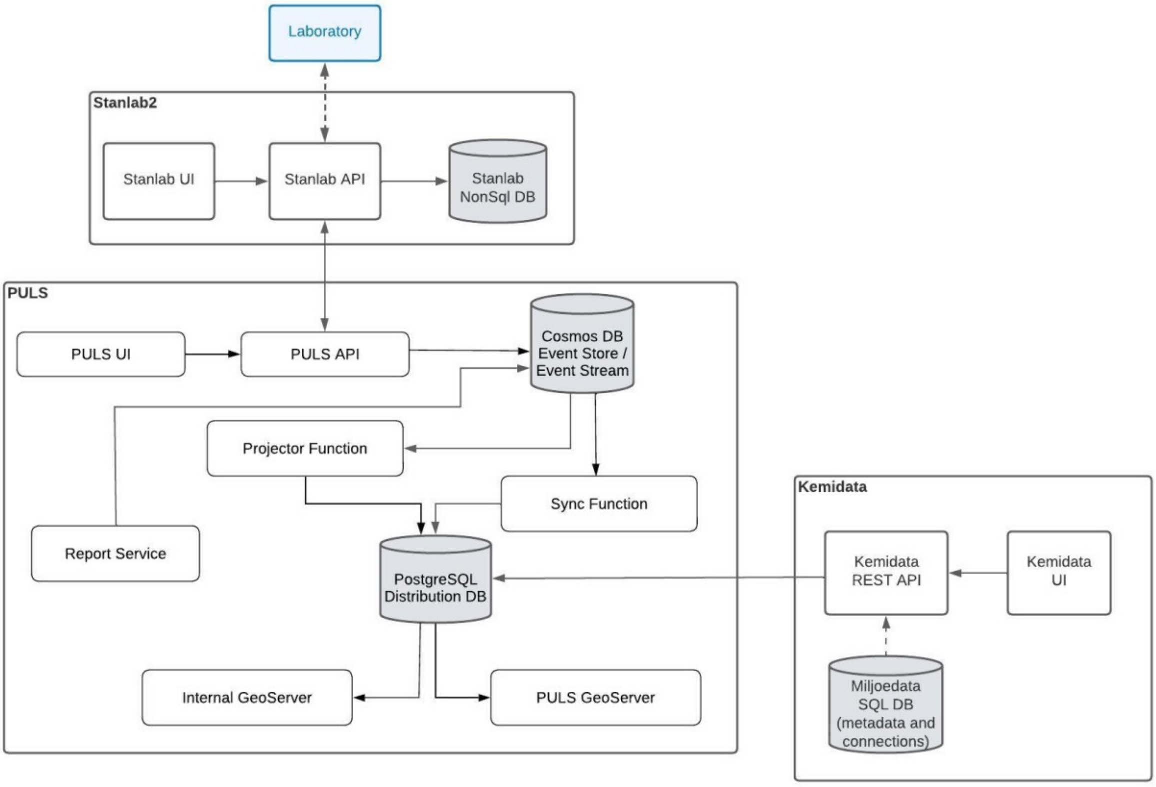

Figure 13: Puls system diagram

Components

DMP has developed the components Emnekatalog, Parameterlisten, Regelkatalog and the Map software component. Further, DMP uses the component Stancodes to share master data with other public offices.

Emnekatalog holds observable species and chemical substances and their measurable unit and methods for research topics. Subject Catalogue - Overview

Parameterlisten holds the official codes (3013 as of March 2026) used for reporting of chemistry, physical and microbiological environmental data. The Danish Ministry of Environment is the data owner.

Regelkatalog holds parameters and rules related to urban wastewater. The Danish Ministry of Environment is the data owner.

Datakatalog holds collections of geographical information and their metadata.

Stancodes holds code lists that are used across public offices. Stancode Lists - Overview

Local lists are code lists that are used by VanDa only. The VanDa solution is described later in this document.

The Map software component makes geographic information discoverable and is used for visualisation.

Data sources

The table below displays an overview of the water environment data.

Raw production data from continuous samples, discrete samples, and results from laboratory	• Groundwater • Drinking water • Surface water including marine water • Urban wastewater • Bathing water • Chemistry data • Biological data (species)
Reference data	• Code lists • Parameters • Rules
Distribution data	• Environmental impact assessments • Data collections for display • Area data • Flood warnings • Validated, enriched water and chemistry data • Geographical information meta data

Data quality

The data quality processes were optimised to a 3-step approach:

1. Automated validation of data syntax and business rules defined in the component Emnekatalog
2. Manual validation by data reporters through secured web site
3. Manual validation by experts at DCE through database Oda Quality Assurance - Overview

Target architecture (DRAFT)

- Front-end interfaces (web, mobile, sensors, open API etc.)
- Load balancer, fire wall, IAM services
- Back-end components
- Shared software components (map, rule engine)
- Standardised data validation
- Low frequency database for manually discrete sampling
- High frequency database for continuedly sampling with sensors
- Distribution of data (virtual, querying, semantic search, load, etc.)

Potential improvements

- Consolidate data models for Puls and VanDa
- Remove unnecessary features (event handling)
- Standardise code lists, both external code lists (Stancode) and internal code lists (local lists)
- Improve governance model for external code lists
- Simplifies rule components (Emnekatalog, Regelkatalog, Parameterlist)
- Simplification of tech stack
- Harmonise quality processes to a standard process across all DMP solutions
- Standardise sampling process for chemistry data and consolidate into single data model
- AI

Input channels	Web	REST API	IoT sensors	3rd-party integration	
Identity and access management	Identity control	User and resource management	API token		
Reference sources	Code lists	Parameters	Rules		
Data models	Observations and measurements data model				
Output channels	Web	REST API	Static views	Database share	Meta data
Computing	Event management	ETL capabilities	ML capabilities		
Data storage	Non-SQL database	Rational database	Blob		

Table 4: Reference architecture