UNIVERSITY OF TWENTE.

Faculty of engineering technology
Construction management and engineering

An enterprise architecture for a paving process control system aimed at flexibility and systematic data collection

Gerko de Vente

Master thesis

15 July 2024

Academic Supervision

Dr. S.R. Miller – Associate Professor

Ir. Q. Shen – PhD Candidate

Corporate Supervision

Dennis van den Wildenberg – Manager

Ivo de Vries – Deputy Director

Summary

This research for the construction company Van Gelder focuses on developing a real-time process control system for asphalt paving operations to address significant issues identified in the current methodology. The main problems include a lack of real-time data, excessive manual labor, and unstructured data collection, which hinder efficient operation and quality control in asphalt paving.

The research is structured around four main objectives. The first objective is to define the requirements for the process control system, with a focus on real-time data collection, data storage and retrieval, integration with existing systems, user accessibility, alerting mechanisms, data security, scalability, reliability, and data quality.

The second and third objectives are to develop the data infrastructure and design a customizable process control system. Both objectives are represented in the conceptual architecture and the logical architecture. For the conceptual architecture, the goal was to define the services that the system needed to provide, which are: real-time information services, analysis, reporting access, and robust data storage management. In the next phase, the logical architecture was drafted. This phase also includes designing a comprehensive data ontology incorporating static, quasi-static, and dynamic data classes to ensure flexibility and customization. The third objective, to design a customizable process control system, continued, focusing on cloud computing for real-time processing and notifications, temporary and redundant data storage, and efficient data transport via 4G. This design ensures that the system can be adapted to the different project requirements and integrates seamlessly into the current enterprise architecture.

The final objective is to future-proof the process control system by anticipating technological advancements and industry changes. This involves strategies for seamless updates, expansions, and integration with emerging technologies, ensuring the system remains relevant and effective over time by looking at the capability gap in the future.

The research concludes that adopting the PQi (Process Quality improvement) methodology, enhanced with real-time capabilities and machine-to-machine communication, is essential for improving the asphalt paving process. This approach aligns with ASPARi's goals of reducing variance in key parameters, improving process control, and enhancing product quality.

Overall, the research aims to create a robust, scalable, and flexible process control system that significantly improves the efficiency and quality of asphalt paving operations at Van Gelder. By leveraging real-time data and advanced digital frameworks, the proposed system seeks to overcome current limitations and set a new standard in road construction practices.

Samenvatting

Dit onderzoek voor het aanemingsmaatschappij van Van Gelder richt zich op het ontwikkelen van een realtime procesbesturingssysteem voor de asfaltverwerking om belangrijke problemen in de huidige methodologie aan te pakken. De belangrijkste problemen zijn het gebrek aan realtime data, buitensporig handmatig werk en ongestructureerde gegevensverzameling, die zorgen voor inefficiënte werkmethode en kwaliteitscontrole van het verwekingsproces.

Het onderzoek is opgebouwd rond vier hoofddoelstellingen. Het eerste doel is het definiëren van de vereisten voor het procesbesturingssysteem, met de nadruk op het realtime verzamelen van gegevens, het opslaan en ophalen van gegevens, integratie met bestaande systemen, gebruikerstoegankelijkheid, waarschuwingsmechanismen, gegevensbeveiliging, schaalbaarheid, betrouwbaarheid en gegevenskwaliteit. .

De tweede en derde doelstelling zijn het ontwikkelen van de data-infrastructuur en het ontwerpen van een aanpasbaar procescontrolesysteem. Beide doelstellingen zijn vertegenwoordigd in de conceptuele architectuur en de logische architectuur. Voor de conceptuele architectuur was het doel om de diensten te definiëren die het systeem moest bieden, namelijk: realtime informatiediensten, analyse, rapportagetoegang en robuust beheer van gegevensopslag. In de volgende fase werd de logische architectuur opgesteld. Deze fase omvat ook het ontwerpen van een uitgebreide data-ontologie waarin statische, quasi-statische en dynamische dataklassen zijn opgenomen om flexibiliteit en maatwerk te garanderen. De derde doelstelling, het ontwerpen van een aanpasbaar procesbesturingssysteem, werd voortgezet, waarbij de nadruk lag op cloud computing voor realtime verwerking en meldingen, tijdelijke en redundante gegevensopslag en efficiënt gegevenstransport via 4G. Dit ontwerp zorgt ervoor dat het systeem kan worden aangepast aan de verschillende projectvereisten en naadloos kan worden geïntegreerd in de huidige bedrijfsarchitectuur.

Het uiteindelijke doel is om het procesbesturingssysteem toekomstbestendig te maken door te anticiperen op technologische vooruitgang en veranderingen in de sector. Dit omvat strategieën voor naadloze updates, uitbreidingen en integratie met opkomende technologieën, waardoor het systeem in de loop van de tijd relevant en effectief blijft door te kijken naar de capaciteitskloof in de toekomst.

Het voorstel concludeert dat de adoptie van de PQi (Process Quality improvement) methodologie, verbeterd met realtime mogelijkheden en machine-to-machine communicatie, essentieel is voor het verbeteren van het asfaltverwerkingsproces. Deze benadering komt overeen met de doelen van ASPARi om variatie in belangrijke parameters te verminderen, procesbeheersing te verbeteren en zo productkwaliteit te verhogen.

Het algehele doel van dit onderzoek is om een robuust, schaalbaar en flexibel procesbesturingssysteem te creëren dat de efficiëntie en kwaliteit van asfaltverwerking bij Van Gelder aanzienlijk verbetert. Door gebruik te maken van realtime gegevens en gea-

vanceerde digitale frameworks, streeft het voorgestelde systeem ernaar om de huidige beperkingen te overwinnen en een nieuwe standaard te zetten in de wegenbouw.

Contents

List of Figures

List of Tables

Acronyms

AI Artificial Intelligence. 90

AO Asphalt Organisation. 10, 23

API Application Programming Interface. 28, 64, 67, 89

AR Augmented Reality. 83, 84, 88

ASPARi ASphalt PAving Research and innovation. 3, 4

ATs Asphalt Teams. 23, 25, 49, 61, 62

BI Business Intelligence. 67, 71

CCP Compaction Contour Plot. 4, 60

CPM Compaction Priority Map. 5, 6, 46, 47, 60, 67

DCAR Decision Centric Architecture Review. 78

DI Data Information. 51

DLT Distributed Ledger Technology. 83

EA Enterprise Architecture. iv, vi, 20–23, 30, 31, 40, 41

EA4DT Enterprise Architecture for Digital Transformation. vi, 16–18, 20–22, 26, 40, 41, 45, 46, 73, 87

ECR Effective Compaction Rate. 5

FTE FullTime-Equivalent. 89

HMA Hot Mix Asphalt. 2

HSE Health, Safety, and Environment. 78

IoT Internet of Things. 4, 65, 66, 84, 88

KPIs key performance indicators. 49

LCV Life-Cycle Value. 80

LoRaWAN Long Range Wide Area Network. 65, 66

LTE Long Term Evolution. 66

MDM Master Data Management. v, 32, 70, 71, 73–75, 77

NB Narrow Band. 66

OSS Operator Support Systems. 5, 12

PAN Personal Area Network. 65

PQi Process Quality improvement. iv, vi, viii, 4, 5, 11, 13, 14, 17, 20–26, 30–33, 36, 38–40, 42, 57, 70, 73, 86, 87

QA Quality Assurance. 2, 5

QC Quality Control. 2, 5

QD Quality Department. 23, 25, 26, 30

SLA service-level agreement. 62

TCP Temperature Contour Plot. 4, 46, 47, 60, 67

WLAN Wireless Local Area Network. 65

WMA Warm Mix Asphalt. 2

WSNs Advanced Wireless Sensor Networks. 83

1
Introduction - Problem context

The Dutch road network is one of the densest in the world, of which more than 130,000 kilometers of paved road [4]. For this, Hot Mix Asphalt (HMA) and Warm Mix Asphalt (WMA) are the most favored road materials in the Netherlands. The quality of these roads depends on the asphalt density and homogeneity [3]. Currently, Quality Control (QC) and Quality Assurance (QA) of HMA and WMA mainly consist of drilling cores at random or selected spots on the road, which will be sent to a lab to determine the density and check if this is within boundaries. This conventional sampling method has several drawbacks [5, 6]; (1) The quality check is limited to the sample points; it is thus not able to determine the uniformity of the entire road. (2) Drilling cores introduce damage to the freshly laid road. (3) This process of drilling and analyzing is time-consuming. The latter has an even more significant additional disadvantage. If the drilled cores do not meet the density requirements, there is no way to achieve additional compaction afterward. In this case, the client might ask for a settlement or a completely new road. A fundamental assumption in the literature is that the compaction temperature is a crucial determinant of asphalt quality [3, 7, 8, 9, 10]. The temperature at which the asphalt is compacted has an effect on the density and mechanical properties of the asphalt. Bijleveld shows in his report that a mechanical property fracture energy could lower up to 30% if compaction takes place outside the temperature window [10]. Figure 1.1 shows the compaction time window based on the temperature window of the asphalt layer. The temperature window depends on the type of asphalt and the asphalt mixture. Logically, the compaction window of HMA is smaller because it reaches its lower-limit compaction temperature faster since the initial temperature is lower.
Chowdhury and Button show that the desired density could still be achieved with temperatures 22 degrees Celsius lower compared to HMA [11]. Bijleveld shows that if this is the case, the compaction window of WMA is longer than HMA [10]. This is beneficial since there is more time for the operator to achieve the desired density. Compaction outside the temperature window can cause over- or under-stressed asphalt [3, 7]. It is thus essential for the roller operator to know the temperature window and the number of passes for the roller to achieve optimal density before the project. During the project, the roller operator should know the real-time temperatures of the asphalt mat and the amount of passes at a certain location. This highlights the importance of the compaction phase during construction and the real-time monitoring of the compaction process. Compaction is a delicate process that is influenced by several causal factors, such as design, execution, logistics, and environmental factors. A change in any of these

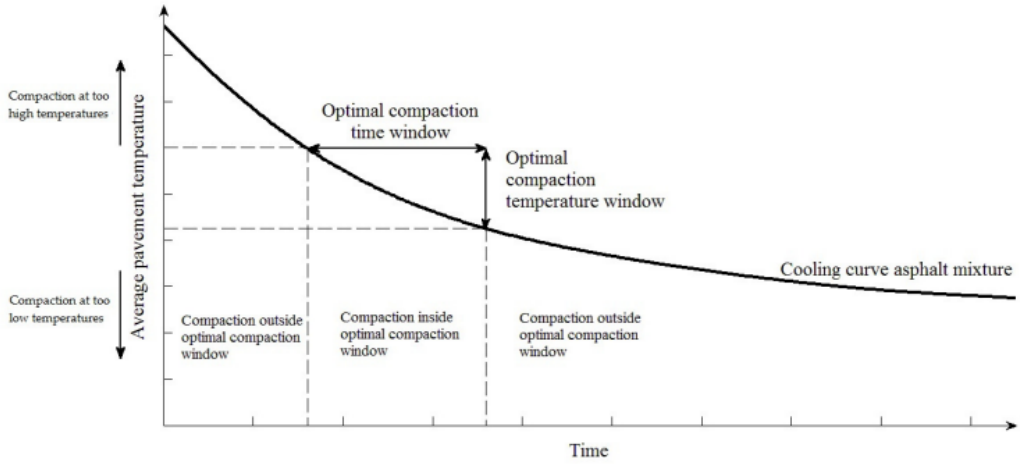


Figure 1.1: optimum temperature window for compaction

factors might cause under- or over-compaction [3]. In conventional asphalt construction practices, there is a lack of real-time information on these aforementioned influential factors. So, there is no possibility of adjusting the strategy on the fly. This leaves the implicit and experience-driven knowledge of the operators to determine the strategy [3].

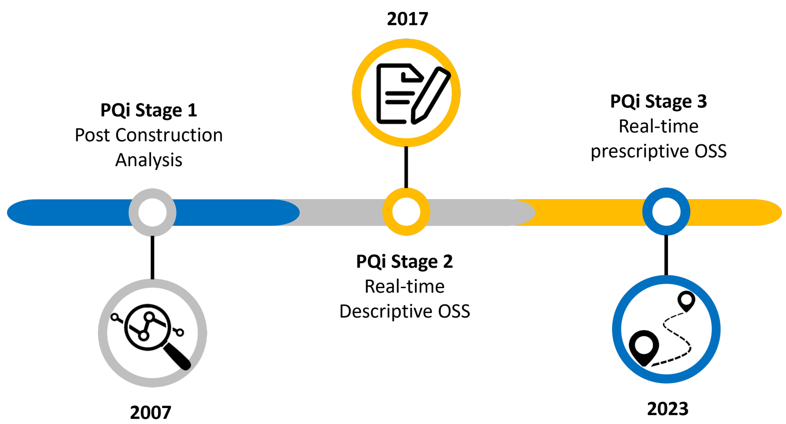


Figure 1.2: PQi Roadmap

Back in 2007 ASphalt PAving Research and innovation (ASPARi) knowledge network was established, see figure 1.2. This Dutch knowledge network consists of the University of Twente, ten large-scale contractors, and the national infrastructure authority [12]. Their first objective was to get a holistic view of the paving process. From this, ASPARi has developed a framework for contractors on how sensors could be used to monitor the asphalt process and gather more explicit knowledge. The main goal of ASPARi is to improve the asphalt paving and compaction process by (1) Developing insights into the asphalt paving process and providing feedback to operators. (2) Reduce variance in key parameters, improve process control, and continuously improve productivity. (3)

Reducing risks for contractors by improving product quality and value for customers [13]. Currently, ASPARi is working with a methodology called PQi, short for *Process Quality improvement*. This method is designed for the improvement of the asphalt paving and compaction process by systematically providing insight into the process, thus making the operational behavior explicit [1]. This framework is based on gathering location data of the paver(s) and compactor(s), real-time monitoring of the surface - and core temperature(s), systematically checking the asphalt density, and monitoring the environmental data, like ambient temperature. This methodology will be referred to as PQi stage 1, see figure 1.2.
The PQi methodology is, in essence, a 5-phase process cycle, which is illustrated in figure 1.3. The cycle starts with the preparation phase, which takes place before construction and includes a meeting with the asphalt team to check the construction site. This is followed by the measurement phase. During this phase, all the essential data (described above) is collected. This data needs to be analyzed and converted into visual data, like Temperature Contour Plots (TCPs) and Compaction Contour Plots (CCPs). The results of the data analysis are then shared with the asphalt team. The lessons learned from such meetings will lead to work improvement [1]. This cycle is thus made to collect, analyze, and share the data to get a better understanding of the asphalt process, which will lead to better asphalt quality.

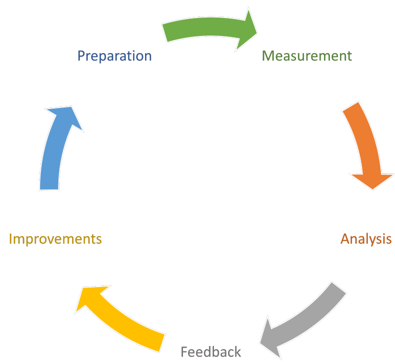


Figure 1.3: PQi cycle [1]

A major limitation of the PQi stage 1 methodology is that the analysis is done after the construction of the road. Although the PQi process helps in getting more insight into the paving process, there is no possibility of adapting to changes or mistakes in real time. This limitation was also highlighted by Makarov. In order to solve this problem, Makarov has provided a framework for a real-time pavement operation support system using machine-to-machine communication. The article demonstrates the use of the Internet of Things (IoT) to connect the individual sensors between different machines to provide real-time information to the roller operator. Case studies showed promising results, where this system was able to provide the operator with the correct information during road construction [7, 3]. This improved PQi methodology will be referred to as

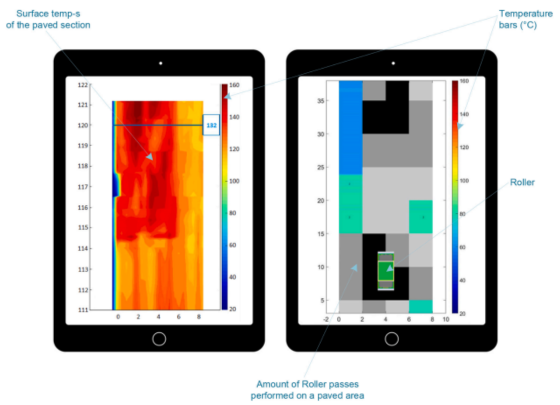
PQi stage 2, see figure 1.2.

Makarov distinguishes between two types of Operator Support Systems (OSS), namely descriptive and prescriptive. Descriptive OSS provides data that shows the current state of the asphalt mat. It, for example, shows the temperature of the surface layer behind the paver, compaction passes, or weather conditions on-site; see figure 1.4a. This leaves the operators with additional information and the need to make their own compaction strategy. Prescriptive OSS, on the other hand, are more guiding. With a holistic view of the compaction process, a prescriptive OSS can create a compaction strategy for the operators in the form of a Compaction Priority Map (CPM), see figure 1.4b. The use of Descriptive systems requires more cognitive workload, which might even have a negative effect on the road quality [3]. The use of prescriptive sounds more favorable since a compaction strategy is made for the operator(s). Results show that the use of descriptive - and prescriptive OSS are beneficial to the Effective Compaction Rate (ECR), see figure 1.5 [3]. The Effective Compaction Rate can be used as a quality indicator for the asphalt compaction only accounting for the cells that have been compacted in the desired temperature window, see equation 1.1 [3]. The closer $ECR_{p,k}$ is to 1, the better the compaction.

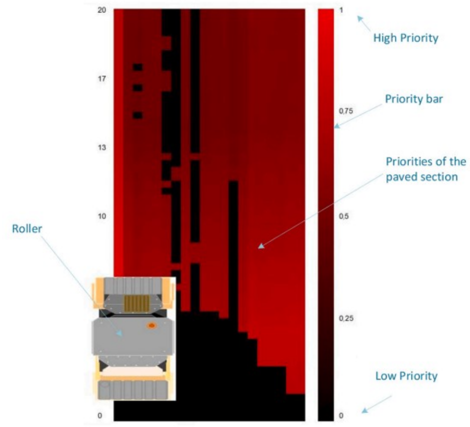
$$\mathrm{ECR}_{p,k} = \frac{N_{p,k}}{N} \tag{1.1}$$

where:
np,k: the number of cells that have been compacted exactly for the target compaction passes ± k, and at least p% of the compaction passes were within the compaction window N: the total number of cells

The case study results show a 67.7% and 115% increase in the ECR for descriptive - and prescriptive OSS relatively compared to no use of OSS [3].
In recent years, the concept of Intelligent Compaction (IC) has gained more momentum [7, 14]. The concept of intelligent compaction has been around for over two decades. IC can be defined as a tool that consists of sensors collecting data and analyzing this data in near real-time to give more insight into the compaction process. Such systems have also proven to increase the efficiency of compaction [3, 7]. This improved PQi methodology will be referred to as PQi stage 3; see figure 1.2.
nowadays many manufacturers have equipped their machines with sensors or provide separate sensors to be mounted on machines, see table 1.1 [15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25]. These systems all aim to provide real-time information on the construction process, such as asphalt mat surface temperature or the number of roller passes. This information can be used for analysis to give insight into the process, act as QC/QA for the client, or can be used for future risk analysis. Makarov has shown a sound proof of concept in the papers that a real-time OSS helps improve the effective compaction rate [3]. Furthermore, there are plenty of market-ready sensor solutions on the market[15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25]. Despite this, there has not (yet) been implemented an operational system likewise on the Dutch road construction market.



(a) Descriptive compaction data



(b) Compaction Priority Map (CPM)

Figure 1.4: fig: Descriptive vs. Prescriptive

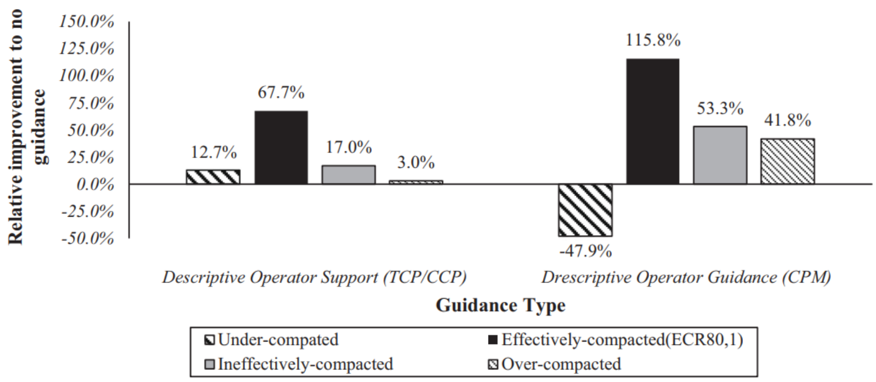


Figure 1.5: ECR80% for descriptive and prescriptive OSS

Table 1.1: Market Ready Solution Overview

Company	Solution name	Solution domain	Solution features
Ammann	ACE-Pro[26]	Compaction	Measures dynamic ground-bearing capacity (kB value), Adjusts frequency and amplitude accordingly
Bomag (Fayat group)	Bomap[17]	Compaction	Surface Covering Dynamic Compaction Control (CCC) provides a clear overview of compaction progress
Caterpillar	Compaction control technologies[25]	Compaction	Indications on the soil stiffness, temperatures, and pass counts with communication between rollers. The amplitude of the drum can be automatically adjusted.
Dynapac (Fayat group)	Matmanager[27]	Paving	Data collection with an interface showing information: temperatures, load, speed, location, time, and start/stops.
Hamm (Wirtgen Group)	Smart compaction[15]	Compaction	Measures rigidity, surface temperature, pass count, and weather data. It can apply automated drum control
Leica	Leica icon Compaction[23]/Paving[24]	Compaction & Paving	Roller register compaction meter value and count pass. The paver has auto steer and auto edge control, which can be sent from project data.
Moba	MCA-3000[22]/PAVE-IR[21]/MOBA-MATIC[28]	Compaction & Paving	Roller registers amplitude, vibration frequency, location, temperature, and compaction end. Gives an indication of the amount of compaction progress. The Paver-IR registers a thermal image of the asphalt mat, and the MOBA-MATIC automatically levels the screed thickness.
Q-point	Q Machines - Pave [29]/ Q Machines - Compaction [30]	Compaction & Paving	Roller measures the stiffness, surface temperature, and the necessary passes based on GPS. The paver shows the temperature information of the mix during loading and paving. Systems are able to communicate with each other in real-time.
Sakai	Intelligent Compaction (IC)[31]		Measures stiffness, location, temperature, and roller passes.
Topcon	C-53 intelligent compaction[32]/3D-MC-Asphalt[20]/Thermal mapper[19]	Compaction & Paving	Measures stiffness, temperature and roller pass. The paving system controls the thickness of the asphalt mat. The thermal mapper registers the surface temperature behind the paver and registers location data.
Trimble	CCS900 Compaction Control System[18]	Compaction	Measures roller passes and stiffness for soil compactors.
Völkel	Pave-Navigator [33]/CCC-Navigator [34]	Compaction & Paving	Roller measures location, asphalt temperature, and oscillation/vibration. Paver measures location, loading, hopper temperature, underground temperature, screw conveyor asphalt temperature, and screed temperature. Systems are able to communicate with each other in real-time.
Volvo	Compact Assist with Density Direct[16]	Compaction	Registers roller passes, temperature, density, and stiffness
Vögele	RoadScan[35]	Paving	Surface temperature registration of asphalt mat with location. Hopper temperature and weather data are collected.

2

Problem description

2.1 Introduction - asphalt process

This part of the report describes the current paving method used in a project by Van Gelder, a well-established Dutch road construction company. It serves as a reference for the rest of the report. The following scene does not 100% reflect the process of every asphalt team within Van Gelder or every project within the same team.
The executor of a team of 7 starts the day with a short meeting early in the morning. The executor looks in the Asphalt app of Van Gelder and sees the project details and all the equipment needed for this project. The day before, he had already sent all the equipment to this place, so there was no need to wait. The app also shows more project details, such as the type of asphalt, with the corresponding temperature window and the layer thickness. The team is also able to view these details in their own app on their work phone, but most of them don't bother and wait for the instructions of the executor. In the shack, the team quickly discusses the work that is planned for today. The executor lets his team know that the goal of today is to pave the base layer for the entrance of a residential area and the binding layer of the road perpendicular to the residential area connected by a roundabout. He tells the team that the base layer needs to have a thickness of 6 centimeters and the binding layer needs to be 4 centimeters. The thickness of each layer has been determined by a structural engineer, so it is not up to the executor to make changes to it. The shack of the asphalt team is next to the construction site, so he shows the team how he wants to execute the work on site. The executor tells the team to start with the base layer of the entrance of the residential area. Since the road width is entirely covered by the paver, there is no room for the rollers to start compaction behind the machine until the road widens enough 50 meters ahead. The executor knows this but does not have much choice to do it differently. Meanwhile, the first truck arrived with asphalt from the plant. After a bit of waiting, the truck is waved to the paver, and the truck driver starts to maneuver his truck into position. When the truck has reached the paver the truck driver starts to gradually dump the asphalt in the pavers' bucket. The executor checks the temperature of the asphalt coming from the truck to confirm it is still hot enough. The paver starts rolling forward with a steady speed. The executor checks the thickness of the asphalt behind the paver by measuring the height difference with the curb. He notices that the layer is a bit too thick and instructs the paver operator to lower the beam slightly. Once the paver enters the broader part of the road, the roller passes to start compacting.

With the first passes, the roller operator sees and feels the asphalt sinking under the immense weight of the roller. The roller is equipped with a temperature sensor so the operator knows when he enters an asphalt patch too close to the paver, which is too hot and backs off a bit and waits in his roller on the freshly laid asphalt. The roller operator knows the temperature window of asphalt in which compaction needs to be achieved. A few minutes later a smaller roller enters the road as well and also starts compacting. Together, they move up and down the road behind the paver until the operators see and feel that no more compaction is achievable, which is also indicated by the surface temperature of the road. The roller operators are skilled and are convinced that they have compacted every bit of the road and have done enough passes. During this process, the executor keeps checking the thickness of the base layer. The executor is furthermore busy planning the coming works and arranging transport for the equipment needed. Once the team has finished the work for today, they load up the paver and the compactor. Most of the crew members grab a beer or a soda to end a hot day of hard work and go home. The next day, they follow the same routine.

2.2 Objectives and related problems

All asphalt of Van Gelder falls under the responsibility of the Asphalt Organisation (AO). Its functions are planning asphalt, equipment, and other businesses that do not directly relate to specific projects, such as transport arrangements and asphalt plant requirements. At the beginning of 2023, a new deputy director was appointed. The deputy director has addressed several challenges that need to be overcome to meet the ambition of 100% flawless asphalt. He mentions (1) the planning being ad hoc. (2) Inefficient asphalt transport. (3) Data processing is done manually(Excel culture). (4) No process control during paving due to the absence of data. (5) Paving and compaction material not always ready for use. Van Gelder has to overcome all of these individual problem areas in order to even come close to the ambition of 100% flawless asphalt. The chapter above describes the current asphalt process within Van Gelder. The operational manager of the AO states that this process is outdated. He emphasizes that there are technologies available on the market to address this. The operational manager addresses several problems with the current way of working, which are described in the chapter above. Although mentioned in the scenario, there is not always a briefing prior to the project. Some team members are thus unaware of the type of asphalt used. Nor does the team reflect on previous projects. He also mentions that the roller operator is, in most cases, not even aware of the compaction window. The deputy director adds that there is no information on the asphalt temperature in the trucks. They have seen bad quality results caused by asphalt that had been too cold when poured into the hopper. The operational manager estimates that 10-12% of the validation samples (drilled cores) do not meet the requirements set in the contract. In most cases, a settlement is made with a certain discount per ton. But in the worst case, a section of the road needs to be removed and re-paved. It is clear that both cases result in failure costs, with the latter being almost a certain loss on the projects. The deputy director and the manager both address the need for more insight during the asphalt process. The deputy director has the specific objective to obtain more data during the asphalt process and use this data to see trends and make predictions for future projects. The operational Manager would like to have the data on-site in real-time and be able to intervene in the process to resolve and prevent mistakes, thus yielding a greater profit. In addition to this, he wants to deliver the projects based on the collected data to reduce and eventually eliminate the need for core sample analysis for every project delivery.

Company objectives

- generate trends and make predictions
- real-time process information
- data-based project delivery

Both the director and manager of the Asphalt division of Van Gelder have highlighted the main problems and their individual objectives. The overall goal is to reduce failure costs by improving the overall asphalt quality. They both have expressed their wishes for a (real-time) process control system for their paving operations. For the past year, the company has partnered with Bomap to collect their roller data and Witos for the paver data in order to collect data and get a better understanding of the paving operations. A methodology was used that was based on the PQi stage 1 methodology to collect this data together with some additional sensors collecting relevant data (weather data and core asphalt temperature). The results of this Van Gelder methodology (PQi stage 1) did not result in achieving either objective. An internal research done prior to this Master's thesis concluded the following problems with the current methodology. (1) Lack of real-time data, which mainly has to do with the manufacturers of the registration systems in general, not only Witos and Bomap. Both platforms were not able to communicate with each other. Additionally, the data could be visualized in real time, but the data itself was only available after the completion of the project. (2) Too much manual labor. This problem is comprised of 3 underlying issues. (2a) Volatile for last-minute changes. This partly had to do with the fact that just one set was equipped with the registration systems. It would most likely still be difficult to orchestrate due to the complex nature of the construction industry, which is prone to sudden changes due to project delays or bad weather conditions. This results in last-minute equipment changes or an entire reschedule of the project. It takes quite some time to properly prepare a measurement following the current Van Gelder methodology, so this resulted in much lost time. (2b) Measurement supervision: As described above, the preparation takes a considerable amount of time, but also, during the data collection, a dedicated crew member is needed to oversee if the data collection is going as planned and should be able to fix any issues arising during the project regarding the measurements. (2c) Unstructured data collection. This has to do with the fragmented data collection. There was no central database where the data was stored, so someone needed to collect all the files, store them locally on a computer, and run the algorithm. It was also very difficult for someone else to get the results since the computer needed special software to run the algorithm, and the data needed to be specifically stored.

PQi stage 1 shortcomings

1. lack of real-time data
2. too much manual labor
   (a) volatile for last-minute changes
   (b) Measurement Supervising
   (c) unstructured data collection

As mentioned in the introduction the framework of Makarov (PQi stage 1 methodology) provides a solution to overcome the problems listed above. Thus far, road construction companies have not adopted such a complete system with machine-to-machine communication. As aforementioned, many machine manufacturers have equipped their machines with their sensors and their corresponding portal[15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25]. The problem, however, is that they either have a solution for one type of machine (paver or roller), or there is no communication between the different types of machines. In many road construction companies, different brands of machines are used. It might be obvious that different manufacturers are not open to connecting their products with other manufacturers. This makes the data storage fragmented. Furthermore, a contractor is, in most cases, not willing to choose just one manufacturer since this makes them dependent on future changes.

As of yet, there are only two ready-to-use solutions that make use of something similar to the provided framework of Makarov [36, 37]. These systems are not dependent on the machine brand. The sensors can be mounted on a wide variety of machines and machine types. These systems can provide real-time information to operators and other essential crew members on the construction site. There remain, however, two major problems, which seem to result in the slow adoption of OSS'. The first problem is with the market ready-to-use solutions, which lack the ability for customization, i.e., open and interoperable systems. The two aforementioned systems, all systems for that matter, make use of descriptive data rather than the preferred prescriptive data suggested by Makarov. The contractor is thus limited by the development rate of the manufacturer of the implemented system. More examples of this are the restricted abilities of the contractor to include additional sensors that they deem important or make changes to the interface. The second problem is with the proof of concept, which does give the ability for customization but lacks an enterprise architecture of the implementation of such a system for systematic inter- and intra-data collection, e.g., inter-project data collection is the collection of data for the same project team but for different projects (on different days). The intra-data collection is the collection of data for multiple projects running at the same time.
An enterprise architecture is essential for the implementation of a process control system. An enterprise architecture can be defined as follows:

> *Enterprise architecture is the process by which an organization standardizes and organizes IT infrastructure to align with its business vision and strategic goals.*

Besides the fact that there are no ready-to-use solutions on the market, the current PQi methodology only provides insight on how the data should/could be collected and how this needs to be presented to the operator [3, 7]. It does not include the human interaction with such a system on a business level, i.e., data appropriate for all operational levels. This is an important starting point for making enterprise architecture, business architecture, and application architecture. The business architecture captures how a business operates. It, for example, includes the organizational structure [38]. The application architecture is based on the business architecture and deals with the user interface and workflows [38]. This problem observation together with the desire of Van Gelder to have a real-time process control system has led to this research.

PQi stage 2 & 3 shortcomings

1. No complete enterprise architecture
2. lack of ability for customization
3. systematic data collection
   (a) Inter-project
   (b) Intra-project

3
Research design

In this section, the research design is formulated to address the problem statement, the detailed research strategy with the methodology followed by the usefulness of the intended research.

3.1 Research objective

As stated in Chapter 2.2, the aim of Van Gelder is to implement a process control system for all their sets/equipment to help improve asphalt quality. This process control system should be based on the needs of Van Gelder and the PQi methodology. The objective of this research is to establish an enterprise architecture for a process control system that still allows Van Gelder to have the ability to make changes to the system, for example, by adding functionality or sensors without being dependent on the manufacturer. Furthermore, this research tries to find a possibility of collecting data for multiple projects simultaneously and for interchanging machinery.

3.2 Research question

The research objective described in section 3.1 has resulted in the following research question:

How to establish an Enterprise Architecture framework within a road construction company to accommodate a real-time process control system that ensures consistent data collection across diverse projects (both inter and intra projects), facilitates flexibility for customization, and promotes future-proofing in the absence of ready-to-use solutions in the market?

This research question is divided into multiple sub-questions

Research Objective 1 (RO1): Requirement definition This objective aims to define the Requirements for the Process Control System. For this, the essential functional requirements for the real-time process control system in road construction are identified, emphasizing effective monitoring, control, flexibility, and customization. The next step will be to align these requirements with established Enterprise Architecture frameworks to foster a standardized and strategic approach.

(a) What are the critical features and functionalities required for real-time monitoring and control in road construction at the operational level?
(b) What are the requirements to ensure flexibility and the ability to customize?

Research Objective 2 (RO2): Develop a Data Infrastructure:
Determine necessary information, data structuring, and ontology for seamless integration within and between road construction projects. Ensure inter and intra-project flexibility and interoperability with established Enterprise Architecture frameworks for a standardized approach.

(a) What model is needed to structure the data?
(b) What data needs to be collected, and where is it needed (ontology)?
(c) How should the data be structured for interoperability within and between projects (intra-project flexibility) and to ensure flexibility for customization?

Research Objective 3 (RO3): Design a Customizable Process Control System:
To achieve this objective, three tasks need to be carried out. The first thing is to explore customization avenues for the real-time process control system, tailoring it to meet evolving project requirements over time. From this, investigate system architecture flexibility to ensure adaptability to various project scenarios. Finally, the designed system should be aligned with or integrated into an overarching enterprise architecture framework for a cohesive and comprehensive solution.

(a) How can the data be processed to accommodate consistent data collection and keep flexibility for customization
(b) How should the data be stored to accommodate this?
(c) How can the data be transported to accommodate this?
(d) How should the data be managed?

Research Objective 4 (RO4): Ensure Future-Proofing of the Process Control System:
Focus on strategies to future-proof the real-time process control system by anticipating technological advancements and industry changes. Investigate system design for seamless updates, expansions, and integration with emerging technologies while maintaining alignment with the broader Enterprise Architecture framework, ensuring adaptability and compliance with evolving industry standards.

(a) Which strategies can be implemented to future-proof the real-time process control system against technological advancements and industry changes?
(b) How can the system be designed to accommodate updates, expansions, and integration with emerging technologies over time?

3.3 Research methodology

This section outlines the methodology employed in this research, focusing on achieving the predefined research objectives outlined previously. The methodology is grounded in a framework designed to facilitate digital transformation initiatives effectively. Specifically, the EA4DT Architecture Development Cycle serves as our guiding framework.
The EA4DT Architecture Development Cycle, see figure 3.1, proposed by Kassner et al. [39], offers a structured and systematic approach to achieving digital transformation within organizations. Based on industry best practices and comprehensive research, this framework delineates a set of procedures and actions to promote efficient digital evolution. Divided into two levels, the enterprise and the project levels, the EA4DT framework captures both the current business structure and the target structure. The Architecture Vision and Architecture Action Plan set the foundation at the enterprise level. In contrast, the project level provides a detailed blueprint for individual projects, starting with an Architecture Outline and progressing through Conceptual Architecture, Logical Architecture, Physical Architecture, and Architecture Governance [39]. For our purposes, the focus will be on implementing a process control system at Van Gelder. The enterprise level will serve as the basis for this individual project and future related projects.
This section will detail how each component of our research methodology aligns with the overarching objectives of this research. The approach, from data collection to analysis and interpretation, will be guided by the principles and practices inherent in the EA4DT framework. Furthermore, any modifications or adjustments made to the framework to suit the specific context of our study will be explained.
By implementing the EA4DT Architecture Development Cycle, the aim is to establish a robust and systematic approach for the implementation of digital transformation within the scope of the research objectives for this thesis at Van Gelder. This introductory statement lays the groundwork for a comprehensive analysis of the methodology, clarifying the rationale behind the selected approach and emphasizing its significance.

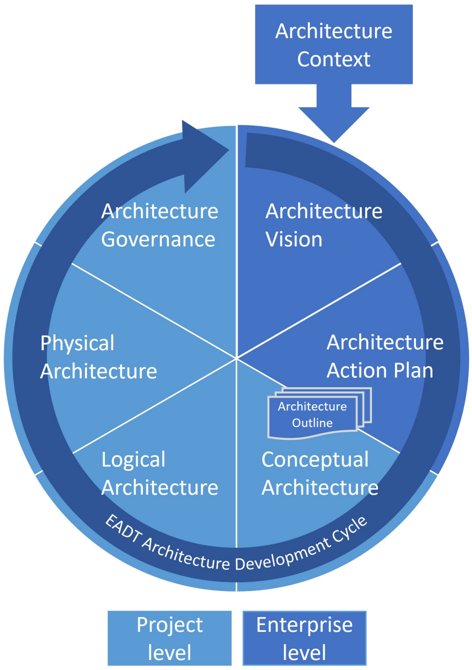


Figure 3.1: EA4DT Cycle

3.3.1 RO1: Define the requirements for the process control system

To achieve the first research objective, the methodology is focused on building a solid foundation at the enterprise level before diving into project-specific information. The process starts with a general analysis of the organization's architectural context and relates the formulated problem to business goals and objectives. From this, the vision is created, which includes an overview of all the principles related to the implementation of the PQi cycle. Here, priority is given to engaging with experts in the form of discussions, where experts and stakeholders directly involved with road construction and digital transformation initiatives can provide valuable insights into the requirements and concerns related to the PQi implementation. The architecture vision is then synthesized to create an architecture plan, which outlines a strategic action plan to meet the objectives of the PQi implementation, including an operating model, a value stream with related capabilities, and a base - and target architecture of the business level. The architecture action plan defines the high-level action plan for the implementation of the PQi cycle. It identifies the architecture constraints and guardrails for the remainder of the projects. Finally, an architecture outline serves as the system's initial blueprint. This outline will explain the chosen project and list all the project requirements that will make up the process control system before defining the scope of the selected project. From this, the conceptual architecture will start the project level in the EA4DT framework. The conceptual architecture stage will translate all the requirements into services for the process control system. Finally, these services will be included in the new target architecture.

3.3.2 RO2: Develop a data infrastructure

The synthesis of findings from the literature review and expert input will inform the development of a data ontology, outlining the necessary information and data structuring requirements to integrate data within and between road construction projects effectively. Emphasis will be placed on ensuring inter and intra-project flexibility and interoperability with established Enterprise Architecture frameworks, fostering a standardized approach to data management in road construction.
This methodology provides a structured approach to developing a data infrastructure aligned with the objectives of RO2, leveraging insights from the EA4DT architecture development cycle to guide the process. Personnel who use or will use process control systems will adapt these methods to the specific needs of road construction projects, using their expertise.
Synthesizing the literature review results and expert contributions will form the basis for developing a data ontology that describes the information requirements and data structure needed for effective data integration within and between research projects. Emphasis will be placed on ensuring flexibility and interoperability between and within projects with established enterprise architecture frameworks to promote a unified road construction data management approach.
This approach provides a structured approach to developing an efficient materials infrastructure aligned with RO2 objectives, leveraging insights from the EA4DT architecture development cycle to guide the process.

3.3.3 RO3: Design a customizable process control system

The third objective, aimed at designing a customizable process control system, is strategically integrated into the logical and physical architecture phases of the EA4DT cycle and forms the backbone of these layers. The focus is on crafting a logical architecture that explores various avenues for customization and flexibility to accommodate diverse project requirements over time. The methodology starts with investigating system architecture flexibility to ensure adaptability to various project scenarios through literature review and case studies (3b, 3c). Aligned with the EA4DT cycle, the logical architecture phase encompasses translating conceptual designs into tangible system specifications. This includes defining system components, interfaces, and interactions, with a particular focus on accommodating customization and flexibility requirements identified earlier in the process.
From this, an in-depth exploration of customization avenues for the real-time process control system, ensuring its adaptability to evolving project requirements. This involves

a comprehensive analysis/summary of project-specific needs and challenges to design a system that can seamlessly scale with the (constantly changing) project dynamics. This will be done through a literature review and expert discussions.
Moving to the physical architecture phase emphasizes selecting technologies and platforms that support the intended logical architecture. This involves evaluating the available options against predefined criteria, focusing on scalability, compatibility, and ease of integration. The results of the integrated logical and physical architecture phases can help develop a customizable process control system that meets the RO3 objectives and can be seamlessly adapted or integrated into the overall enterprise architecture framework. This approach ensures a cohesive and comprehensive solution that meets specific project needs and broader organizational goals. Although hardware acquisition is still in a theoretical phase at this time, the approach continues to envision a scalable and adaptable system architecture to lay the foundation for future implementation. Decision-centric architecture validation will be used to validate the architecture. This validation method includes analyzing the architecture decisions and reviewing them with stakeholders. This method is chosen since it is effective and requires relatively little time and recourses [40].
Finally, a literature review is done to include data management. The study describes the importance of data management and relates the designed architecture to the best data management architecture for this use case. From this, further investigation is done on the implementation of this architecture for data management.

3.3.4 RO4: Ensure future-proofing of the process control system

This phase is dedicated to formulating strategies and the implementation of the necessary measures in order to guarantee the future of the process control system and its ability to adapt to technological advances and industrial changes. The approach includes exploring proactive strategies to anticipate future developments and designing systems to accommodate upgrades, expansion, and seamless integration with new technologies. Future expected requirements are listed, and a literature review will be used to find emerging technology trends related to these requirements to serve as a potential solution. These solutions will be reviewed for usability in future-proofing the process control system. This review serves as a basis for identifying potential risks, opportunities, and challenges that may affect the future viability of the process control system. With this literature analysis, the study explores ways to incorporate technological advances and industrial changes. Throughout the process, a conscious approach that prioritizes adaptability and compliance with ever-changing industry standards is maintained. By combining the research objectives with a structured approach to defining the EA4DT cycle, a systematic and coherent approach was created to guide the development of scalable process control systems tailored to meet the unique requirements of road construction projects.

3.4 Practical and scientific relevance

3.4.1 Practical Relevance

The practical importance of this research lies in its direct impact on the construction industry, especially road construction. This study aims to address the challenges construction companies face when developing systems for the timely control of processes. The results of this research have the potential to revolutionize project management in the construction industry by providing a systematic and adaptable approach to monitoring and control. Implementation of the proposed enterprise architecture framework will bring real benefits to construction companies. It will improve the efficiency of real-time operations, simplify data collection for different projects, and offer customized options for various project needs. The practical application of this research can help reduce costs, shorten project duration, and improve the overall efficiency of road construction project management.

3.4.2 Scientific relevance

This research significantly advances project engineering and construction project management by filling a critical gap in the literature. Comprehensive frameworks exist for data collection and analysis in construction projects, but integrating EA principles is absent. Incorporating EA principles is essential for successfully implementing the proposed Process Quality Improvement (PQi) framework, as EA provides a structured approach to aligning business objectives with technological solutions. By developing a robust, tailored EA framework for road construction projects, this research addresses immediate challenges and lays the groundwork for similar complex future projects. The final EA framework serves as a scaffold, offering a blueprint to integrate process control systems into construction projects, enhancing overall project efficiency and success.

3.4.3 Research scope

Developing an entire EA4DT takes considerable time and resources. Since the time for writing is limited, it is essential to determine the scope clearly and set boundaries to keep the research feasible. As stated in section 3.1, the objective is to develop an EA for a process control system based on the PQi methodology, see figure 1.3. The aim is to complete the entire EA4DT framework as proposed by Rozo [39]. The Enterprise as a whole level will cover the entire PQi cycle. Only the process control system will be covered at the project level. A further limitation is the actual development of this system. To keep the research feasible for the given time planning, only a conceptual system will be developed and validated.

- **Scope**: The scope of this research is focused on developing an enterprise architecture (EA) for a process control system based on the Process Quality Improvement (PQi) methodology, as proposed by Rozo [39]. The EA4DT will encompass the entire PQi cycle at the enterprise level, while only the process control system will be covered at the project level.

- **Limitation**: Due to time constraints, the development of the process control system will be limited to a conceptual level. A detailed implementation of the system is beyond the scope of this research.

The boundaries mentioned above and limitations increase the feasibility of the intended research for the set time frame. However, risks are still attached that can lower the feasibility of succeeding in the desired time frame. These risks are listed as follows:

- **Technical Complexity**: Developing an enterprise architecture for a process control system involves addressing various technical challenges, such as system integration, data management, and scalability. Ensuring the feasibility of addressing these complexities within the research scope is crucial.

- **Resource Constraints**: Limited resources, including time and funding, may challenge the comprehensive development and validation of the proposed enterprise architecture. Mitigating these constraints and optimizing resource utilization will be essential for project success.

- **Data Availability**: Availability and quality of data needed for developing the enterprise architecture and validating its effectiveness may vary. Access to relevant and reliable data sources is essential for conducting accurate analyses and making informed decisions.

These are the 3 main risks associated with this program. One might say that stakeholder engagement should also be mentioned here, but from experience, it can be noted that this variable will not pose any risk for this research. The most significant risk for this research will be the resource constraints; money is the most important variable, but Van Gelder is willing to invest in developing a proper process control system. The cost depends on the proposed solution at the end of this research, but the money needed to develop the conceptual system is minimal; thus, it should not pose a serious risk for this research. The time required for the desired outcome poses a more significant risk and is highly related to the technical complexity of the process control system developed during the EA4DT. Suppose the development of such a conceptual system is in the intended time frame. In that case, the researcher might opt to lower the technical complexities of the system by reducing or changing the system's requirements. Finally, the data availability is essential to develop the EA for the conceptual system. The researcher can access most scientific articles through the University of Twente, which should mitigate the data availability risk.

4

Requirement definition for Process control System

This chapter will cover the first research objective. This results in a good overview of all the functional requirements for a real-time monitoring system and all the additional requirements to make such a system flexible and open for customization.
As described in the methodology, the architecture context is first drafted, and next, the architecture vision is created. Both are tailored to the needs of Van Gelder and cover the entire PQi cycle. From this, the action plan and the architecture outline are presented with a focus on the process control system.

4.1 Architecture Context

Van Gelder is visualizing the implementation of the PQi strategy to improve the overall asphalt quality. The objectives of Van Gelder are elaborated in 2.2, in short:

Objectives

- generate trends and make predictions
- real-time process information

This section is the starting point of the EA4DT cycle. It consists of the business rationale that initiated and motivated the need and development of an Enterprise Architecture for Digital Transformation. It starts by defining the scope of the EA from which the strategies and principles are defined. Finally, the business goals, objectives, and requirements are defined. Business goals are achievements that an organization aims to accomplish. Strategies are the plans and approaches that have to be taken in order to accomplish the business goals. The objectives are measurable steps for the organization to take in order to accomplish the business goals. The principles are statements that guide the decision-making process in the organization in order to accomplish the business goals. Finally, the requirements are clear conditions and capabilities that the system needs to meet or have.

4.1.1 Initiative scope

As stated in section 2.2, all asphalt of Van Gelder falls under the responsibility of the AO. As described earlier (Section: 3.1), the main research objective is to improve the asphalt quality with the use of a process control system in the PQi framework. For this reason, the scope of this EA focused on the Quality Department (QD). The QD is part of the AG. The QD is responsible for the asphalt quality by (1) developing and testing asphalt types. (2) providing protocols for paving and informing/educating executing staff on the correct method. (3) validating the asphalt quality by drilling cores, analyzing the cores in the lab, storing the collected data, analyzing the data, and reporting back to specific projects if needed. The latter shows the importance of including the QD in the EA since they are the ones that will use the data to change/improve 1 and 2. The Asphalt Teams (ATs) are not part of the QD but play an essential role in data collection during construction. As a direct cause, they also have to change their way of working based on the technology we want to implement. Also, indirect causes affect how they work since there might be more different types of feedback from the QD. Together, the QD and the ATs cover the entire PQi cycle; see Figure 1.3. This EA thus focuses on the QD and ATs and aims to improve road quality by accommodating a real-time process control system that ensures consistent data collection across diverse projects and facilitates flexibility for customization.

4.1.2 PQi strategy

The PQi strategy is defined by Data-driven culture, data management and governance, and flexible and scalable infrastructure, see figure 4.1. The strategy includes the people, process, and technology to ensure a complete overview of the initiative. For the people in the PQi, it is essential to get a good understanding of the data to understand the process better and be able to make changes accordingly. At this point, the knowledge gap at Van Gelder is high, making this an essential part of the strategies. The process is covered by data management and governance; here, it is essential to have clear roles and responsibilities to manage the data. Besides that, it is also essential to develop policies and standards for storing and making changes to the data if needed. This all has to happen to keep the data structured and reliable.

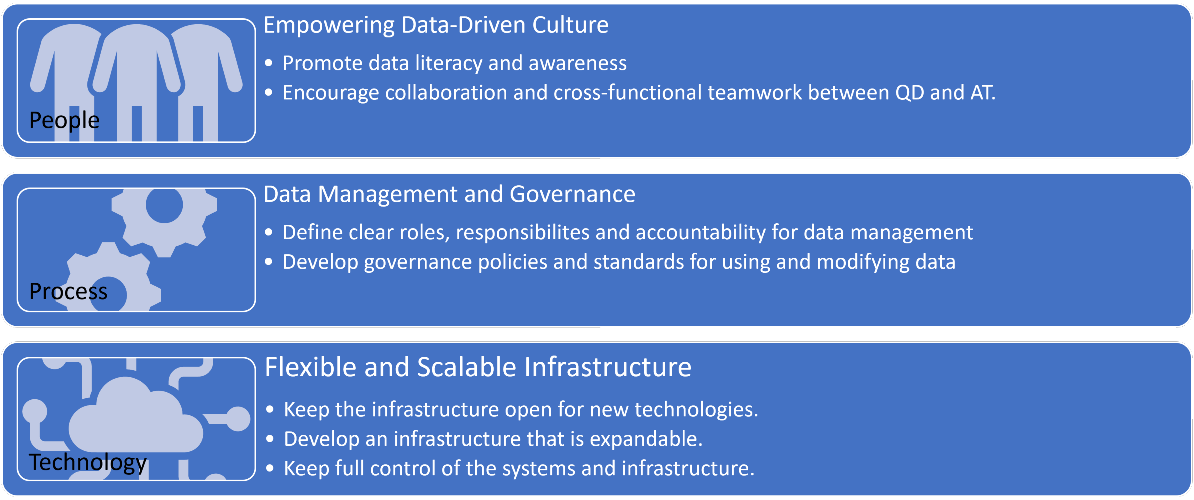


Figure 4.1: Van Gelder strategy for PQi initiative

The technology aspect is focused on flexibility and scalability. To do so, it is essential to keep the infrastructure open to possible new and better technologies that could improve data quality. Furthermore, complete control over the infrastructure/systems should be available to stay flexible and customize the system/infrastructure. Finally, the infrastructure should be able to expand if more data is needed in the future.

4.1.3 Business goals, objectives and requirements

The PQi initiative at van Gelder has several goals, objectives, and requirements that should be achieved by implementing the PQi.

Improve feedback
It is important that lessons learned from process data and asphalt quality are not kept to the people in the lab and managers behind the desk but shared with the people executing the job. The ones that actually pave the roads need to be informed about the results of (all) their projects. This allows people to learn from 'mistakes' or better understand the process, which might help them in their next project. The data shown to the construction crew must be understandable and might need explaining. The QD needs to guide the crew members to help them understand the data and have to listen and act upon their feedback. All this also helps to establish a better relation between the QD and the ATs.

Improve data collecting
The data collection is essential for the PQi cycle. The data is the basis for the PQi cycle and must be high quality and collected from every project. It should also consist of all the essential sensors to collect the data for the entire project duration. With the PQi initiative, the data collection should be more systematic and organized. Furthermore, it should be less intensive by automating parts of the data collection. Another requirement is that the collection of data is systematic. For this, the system should be flexible by working with multiple sets collecting data at the same moment (intra-project) and for projects on different days with different assets (intra-project). Furthermore, the data should be accessible in real-time for interventions and after the project for future analysis.

Improve forecasting
The PQi should help forecast future projects' overall quality based on known variables. This should help to determine if the desired quality is feasible. This can help to get better agreements between the contractor and the client. It can also protect the contractor from executing a job that is likely to fail to meet the density requirements. Still, it also could result in a different set than originally planned to accomplish the job.

Improve effectiveness of suggested improvements
Another goal of the PQi initiative is to improve the effectiveness of suggested improvements. With the PQi cycle, there will be better data available and a more established relation between the QD and the ATs. Both contribute to better-suggested improve-

ments. The suggested improvements should be feasible and tested properly. Due to the PQi cycle, the improvements, either procedural or technical, will show their effectiveness in the analysis phase of the cycle. This might prove whether the implemented change is actually an improvement and whether it is worth the cost.

Quality assurance
The final goal of the PQi initiative is the ability to deliver the project to the client solely based on the data collected during the project. The QD should be able to verify the road quality. This would generate a better overall view of the project quality rather than selected sample cores. In addition to complete quality control, this analysis should take considerably less time than the conventional method.

4.2 Architecture vision

The strategies and objectives have been defined for the PQi initiative in the Architecture Context. The next phase in the EA4DT cycle is the Architecture Vision; see figure 3.1. This phase includes the architecture principles, enterprise capabilities with value stream cross-mapping, The EA operating model, and the baseline and target architecture models [39].

4.2.1 Architecture principles

1. **Business principle**

Name:	Data-Driven Decision Making.
Statement:	Data analysis drives all decision-making processes within the organization.
Rationale:	By basing decisions on data insights, the organization can ensure alignment with strategic objectives and improve operational efficiency.
Implications:	Requires investment in data analytics tools and employee training. Encourages a cultural shift towards data literacy and evidence-based decision-making.

Table 4.1: Business Architecture Principle: Data-Driven Decision Making

Name:	Stakeholder Engagement.
Statement:	Involving stakeholders in decision-making processes and fostering collaboration between business units.
Rationale:	Ensures that architecture decisions align with business needs and priorities. Improves buy-in and support for architecture initiatives.
Implications:	Requires establishing clear communication channels, stakeholder workshops, and continuous feedback mechanisms.

Table 4.2: Business Architecture Principle: Stakeholder Engagement

2. **Data principle**

Name:	Data Quality Management.
Statement:	Ensuring high-quality data through standardized processes and validation mechanisms.
Rationale:	High-quality data is essential for accurate decision-making and reliable insights.
Implications:	Requires implementation of data governance policies, data cleansing procedures, and continuous monitoring of data quality metrics.

Table 4.3: Data Architecture Principle: Data Quality Management

Name:	Master Data Management.
Statement:	Establishing a single, authoritative source of master data to ensure consistency and accuracy across the organization.
Rationale:	Prevents data inconsistencies, redundancies, and errors arising from decentralized data management practices.
Implications:	Requires defining data ownership, data stewardship roles, and implementing data integration and synchronization mechanisms.

Table 4.4: Data Architecture Principle: Master Data Management

3. **Applications principle**

Name:	Modular Application Development.
Statement:	Developing applications as modular components to enhance flexibility, scalability, and maintainability.
Rationale:	Modular architecture facilitates easier integration of new features, updates, and changes.
Implications:	Requires adherence to standardized interfaces, Application Programming Interface (API)s, and development practices. Promotes reuse of components and reduces development time.

Table 4.5: Applications Architecture Principle: Modular Application Development

Name:	Microservices Architecture.
Statement:	Designing applications as a collection of loosely coupled, independently deployable services.
Rationale:	Enables agility, scalability, and resilience in application development and deployment.
Implications:	Requires robust service discovery, API management, and distributed data management capabilities. Promotes autonomy and flexibility in development teams.

Table 4.6: Applications Architecture Principle: Microservices Architecture

4. **Technology principle**

Name:	Cloud-First Strategy.
Statement:	Prioritizing cloud-based technologies for scalability, flexibility, and cost-effectiveness.
Rationale:	Cloud computing offers on-demand resources, scalability, and reduced infrastructure costs compared to on-premises solutions.
Implications:	Requires migration of existing systems to cloud platforms, integration with cloud services, and adoption of cloud-native architectures and practices.

Table 4.7: Technology Architecture Principle: Cloud-First Strategy

Name:	Modularity and Scalability
Statement:	The architecture is designed to be modular and scalable to accommodate future growth and changes in demand.
Rationale:	Modularity enables components to be developed, deployed, and updated independently, while scalability ensures that the architecture can handle the increased workload and user demands without degradation in performance.
Implications:	Adopting microservices architecture, containerization, and cloud computing technologies can enhance modularity and scalability. Additionally, implementing automated scaling mechanisms and monitoring tools is essential for managing scalability effectively.

Table 4.8: Technology Architecture Principle: Modularity and Scalability

Name:	Flexibility for Customization
Statement:	The architecture allows customization to meet specific business requirements and user preferences.
Rationale:	Customization empowers users to tailor applications and workflows to their unique needs, improving user satisfaction and productivity.
Implications:	Providing configurable features, user-friendly customization interfaces, and extensive documentation are essential for enabling customization. Additionally, establishing governance mechanisms to manage customizations and ensure compatibility with future upgrades is necessary to maintain system integrity.

Table 4.9: Technology Architecture Principle: Flexibility for Customization

4.2.2 EA operating model

Intel distinguishes three operating models for an EA: a Centralized, Distributed, and Federated model [41]. The business structure at Van Gelder and the initiative scope would suggest the need for a federated model, where multiple value streams report to different managers with a responsibility towards each other. However, given the size of Van Gelder, a centralized EA model is chosen for the PQi initiative; see figure 4.2. The centralized model is chosen since it allows for simplified control and governance compared to other models. All architects report to a single organization, so there is better alignment with the business strategies [41]. Furthermore, with a centralized model, Van Gelder can respond more quickly to business strategy changes or market conditions [41]. Each domain architect (Business, Data, Application, and Technology) reports to the domain architect, who, together with the project architect, reports to the lead enterprise architect. Both the lead enterprise architect and the project architect are responsible for the overall architecture governance. The lead enterprise architect reports to the QD manager and the region managers. The QD manager reports to the asphalt organization manager, who, together with the region managers, reports to the general director of Van Gelder.

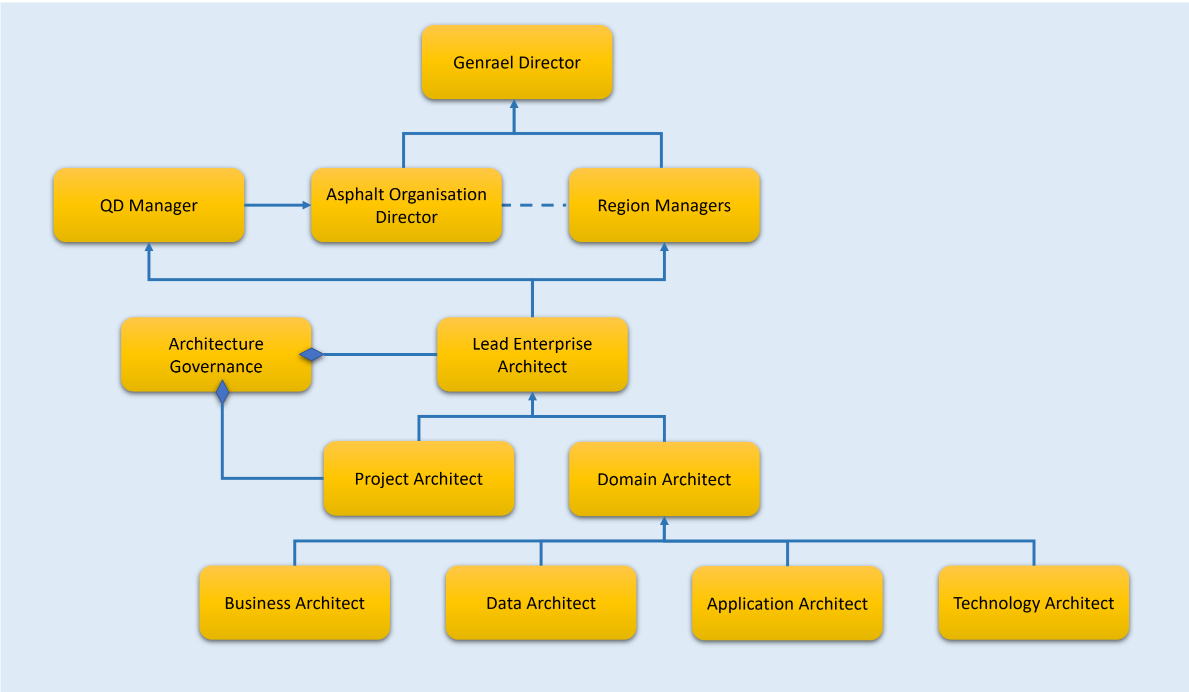


Figure 4.2: EA operating model for PQi

4.2.3 EA capabilities and value stream cross mapping

The improvement of asphalt quality is based on the PQi cycle. For this cycle, a value stream map has been created; see figure 4.3. For each value stream, the necessary capabilities are defined. The first step is acquiring data, which corresponds to the measurement phase of the measurement phase of the PQi cycle. This is followed by organizing the collected data and the data analysis. The data organization value stream ensures that the data is formatted and stored correctly for the next value stream: data analysis. Here, with descriptive analysis, the data can be manually examined. With data mining capability, predictive analysis updates models for what-if scenarios. Both value streams correspond to the analysis phase of the PQi cycle. The feedback value stream is comprised of the visualization of the data and the delivery of the data to its end users. The improvement value stream consists of process optimization, which allows for the design of a more efficient process. The continuous improvement planning should create policies and plans to improve data quality and improve the PQi cycle itself. Quality metrics implementation allows for quantification of the effectiveness of the implemented improvements. Change management is needed to manage a smooth transition when new changes are implemented and to increase the acceptance of the intended/implemented change. This stakeholder engagement is also an essential capability for improving the value stream. This will help create better improvements and increase the acceptance level. Finally, data source identification and data integration planning are both capabilities that are part of the preparation's value stream. The first is responsible for identifying the source needed to acquire the additional, different, or improved data. Data integration planning is, for example, responsible for integrating new/changed data sources into existing infrastructure. Currently, Van Gelder does not possess the envisioned capabilities (as described above). Van Gelder can deliver PQi stage 1 reports and thus possesses the related capabilities to some extent.

Figure 4.3: PQi capabilities with value streams mapping

Table 4.10: PQi Capabilities and definitions

Capability	**Definition**
Data Acquisition	acquisition of information from some physical phenomena. The main process is to sample the signals that convert the sensor's analog value (electrical signal) to a digital one and be manipulated by a computer [42].
Data Mitigation	The process of transferring data from one computing or storage environment to another [39].
Data Quality Assurance	Data Quality Assurance is the systematic process of ensuring that data meets the defined quality standards by implementing measures and procedures to prevent, detect, and correct errors or inconsistencies in data, thereby ensuring its accuracy, completeness, consistency, reliability, and timeliness [43].
Data Transformation	Embodies the process of standardization and mapping of data entities to corresponding target representation upon identification of data errors. [39].
Data Integration	The data management function for consolidation data from multiple sources and ensuring consistency across controlled redundant data with a "golden version" [39].
Data Cleansing	Focuses on activities that correct and enhance values of data elements. Correction activities are pushed to source systems whenever possible [39].
Master Data Management	MDM is the process of defining how master data is created, integrated, and used throughout the enterprise. Master data is the authoritative, most accurate data about key business entities [39].
Descriptive Analysis	Represents the manual examination of data or content characterized by traditional BI and visualization techniques, e.g., pie, bar charts, line graphs, tables, or generated narratives [39].
Predictive Analysis	Is the capability to perform what-if analysis that allows an organization and its users to create and test models based on actual data, then project future results [39].
Data Mining	Refers to a technique in the area of data analytics. It mainly focuses on revealing patterns in data using various algorithms [39].

Continued on next page

Table 4.10: PQi Capabilities and definitions

Capability	Definition
Data Visualizations	A way to represent information graphically, highlighting patterns and trends in data, helping the user to achieve quick insights [39].
Data Delivery	Refers to the optimal point of impact where data is delivered, whether to support human activities, embedded analysis into business processes or support predictive analytic activities [39].
Process Optimization	The ability to analyze and redesign existing processes to improve efficiency, effectiveness, and quality [44].
Continuous Improvement Planning	Developing plans and strategies for continuous improvement of data quality and data analysis processes [45].
Quality Metrics Implementation	Quality metrics implementation refers to the process of incorporating measurable standards. It benchmarks into organizational practices to assess and ensure the quality of products, services, or processes [46].
Change Management	Change Management is the structured approach to transitioning individuals, teams, and organizations from their current state to a desired future state to achieve business objectives effectively [47].
Stakeholder Engagement	Stakeholder engagement refers to involving individuals, groups, or organizations with an interest or stake in a particular project, decision, or initiative. It involves communicating with stakeholders, understanding their concerns, soliciting their input, and incorporating feedback into decision-making processes [48].
Data Source Identification	Data source identification is the process of systematically identifying and cataloging the various sources from which data can be collected or obtained. It involves identifying internal and external sources, understanding their characteristics, and assessing their suitability for fulfilling specific data requirements [49].
Data Integration Planning	Data integration planning involves developing strategies and methodologies for integrating data from diverse sources into a unified environment, facilitating seamless data exchange and interoperability [50].

4.2.4 Baseline architecture

Figure 4.4 shows the baseline business architecture. This represents the current architecture at Van Gelder at the business level. It can be seen that the data (core samples) are collected by a field technician. The lab technician then analyzes the samples. Only when the analysis of a sample has a density outside the accepted spectrum is a peer involved in validating the result. The overall analysis is reported by the head of the quality department and sent to the operational manager, asphalt executor, and asphalt coordinators. Asphalt-type consultations are regular meetings that involve the head of the quality department, all the technicians, and the asphalt developer. These meetings are used to discuss the period's results before the meeting. These meetings are vital since they ensure that the developed asphalt types match the executed projects. Tactical management meetings ensure a smooth running division that aligns with the strategic vision. The asphalt executor is responsible for the process management of the projects on site. A monthly feedback session with the asphalt coordinators, head quality department, asphalt executor, and the operational manager gets people on the same page and helps optimize the current procedures.

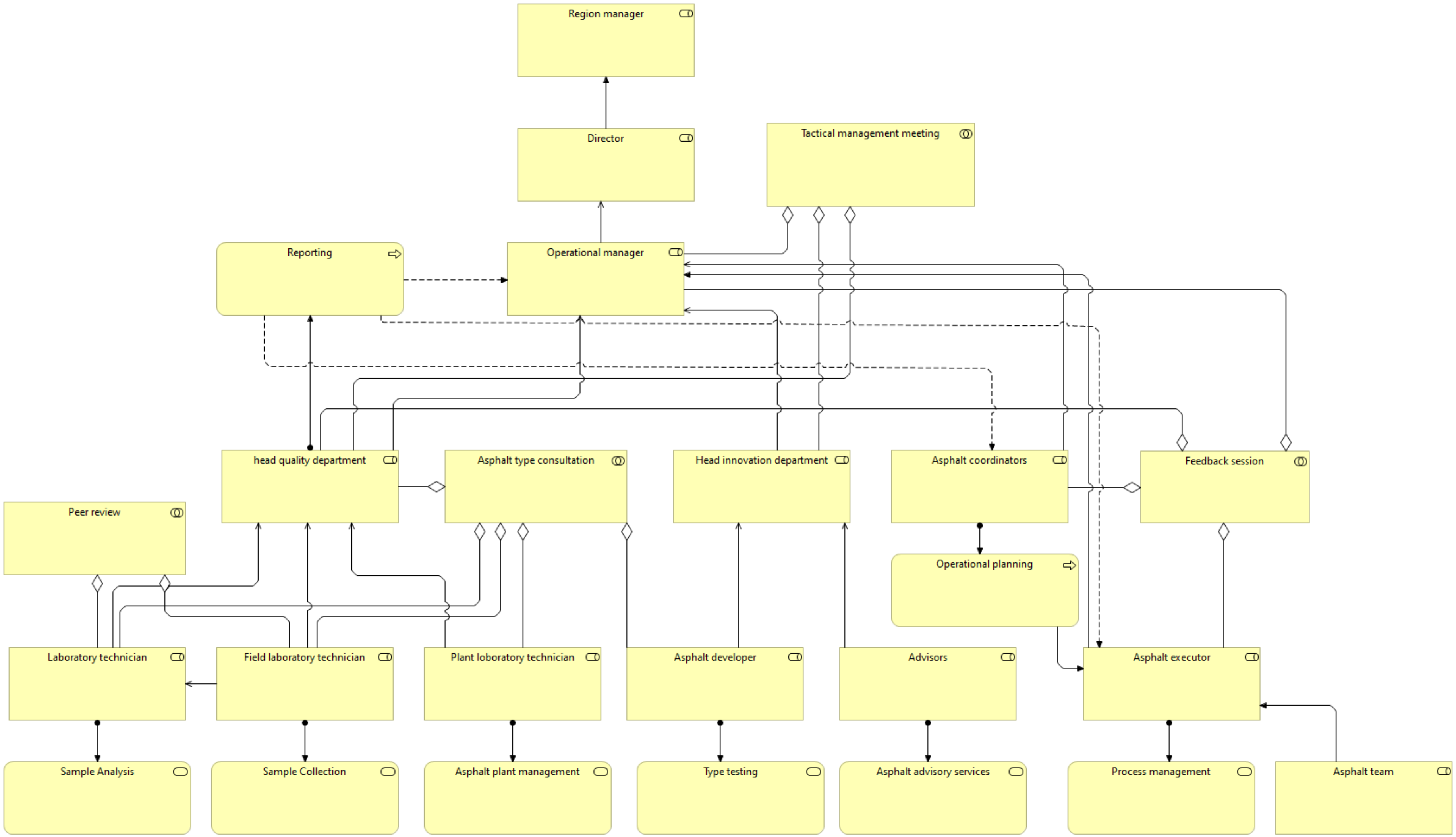


Figure 4.4: Baseline architecture - business level

4.2.5 Target architecture

The target architecture as can be seen in figure 4.5 has the same overall structure as the baseline architecture (figure 4.4). This is because of the complexity and the financial investment to introduce new staff or move people around. To accommodate the introduction of the PQi initiative and its corresponding objectives (section 4.1.3), three new functions have been created. The first new function is a data analyst. This data analyst is responsible for the analysis of all the project data. This, together with the core sample analysis, is reported for internal use. These reports can be generated separately or can be merged into one report. The data analyst is also responsible for the data-based road delivery. For this, the data analyst needs to combine the core samples with the project-specific data and, together with the process manager, validate the control system. In this architecture, the data analyst serves the head innovation department, but this might ultimately change so that the data analyst falls under the supervision of the head quality department or the process manager. The reason for this is that this function will probably be fulfilled by one of the current advisors; this keeps the changes, compared to the baseline architecture, minimal and will probably contribute to a smoother integration of the PQi initiative. The process manager is the second new function added to the architecture. This function is responsible for managing the paving control system and the PQi integration at Van Gelder. The process manager participates in the feedback sessions but will also be on project sites to get direct feedback from the asphalt teams. The process manager falls under the supervision of the operational manager. The last newly added function is the data engineer, who is responsible for the data infrastructure and eventual changes/updates to the system. The data engineer is supervised by the process manager. The data engineer might also be outsourced to another company, but this is highly dependent on the workload's gravity and the implemented system's complexity.

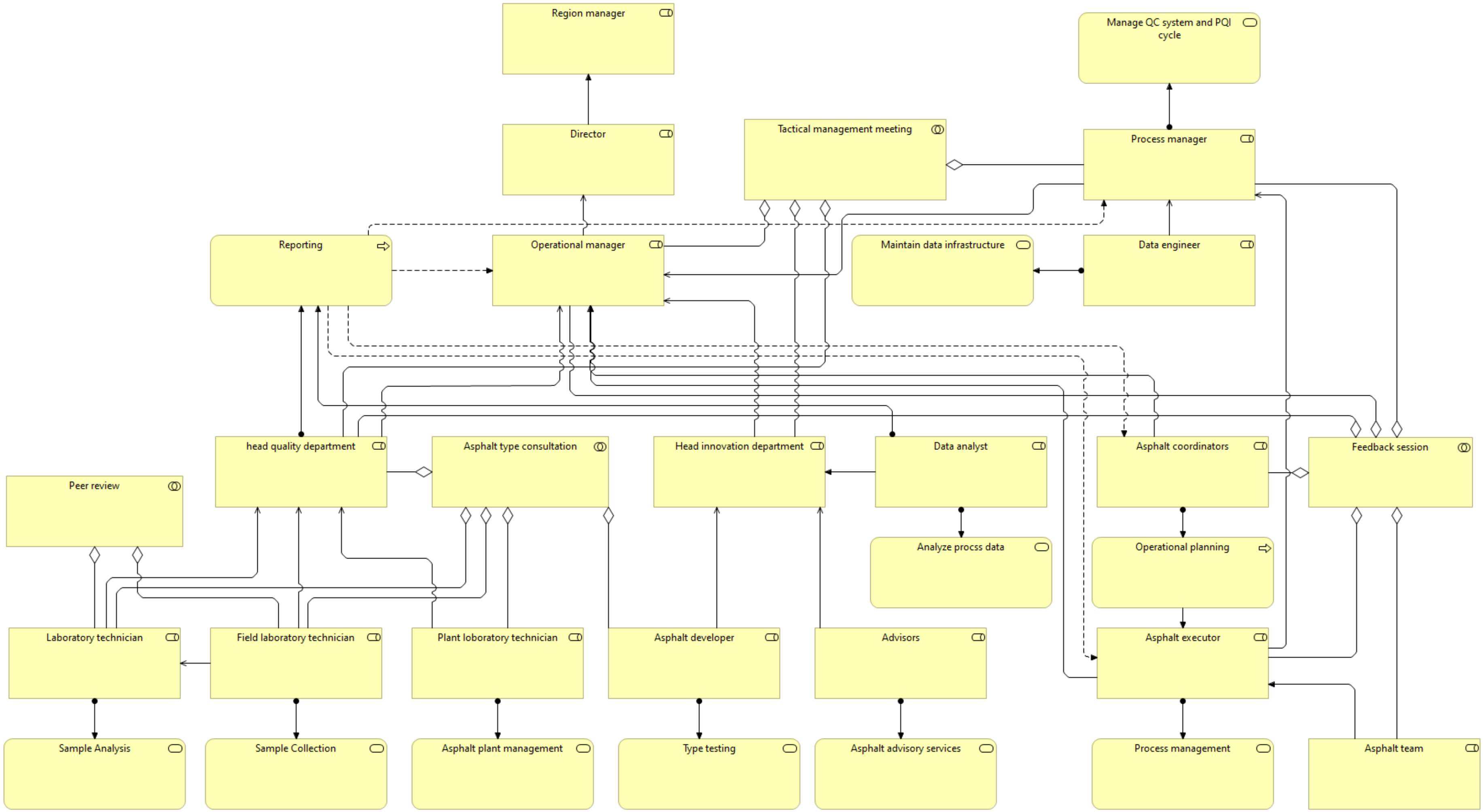


Figure 4.5: Target architecture - business level

4.3 Architecture action plan

4.3.1 High-level action plan

In order to enhance the efficiency and effectiveness of Van Gelder's quality department, several key projects have been identified. Each project plays a crucial role in contributing to the overall system and improving the quality assurance processes within the organization with the implementation of the complete PQi cycle.

Real-time Process Control System
Developing a real-time process control system is essential for optimizing data collection processes and reducing manual effort within Van Gelder's quality department. By implementing automation solutions, data collection processes can be streamlined, resulting in increased efficiency and accuracy. Additionally, integrating real-time data access capabilities enables timely interventions and post-project analysis, allowing for proactive decision-making and improved quality control.

Predictive Analytical Model
Implementing a predictive analytics model based on historical data is paramount for forecasting the overall quality of future projects at Van Gelder. By leveraging past project data, this model can identify patterns and trends, allowing for more accurate predictions of project outcomes. This proactive approach enables Van Gelder to anticipate potential quality issues and take preemptive measures to address them, ultimately enhancing project success rates and customer satisfaction.

Reporting Standardisation
Standardizing reporting formats for road quality assessments ensures consistency and clarity across Van Gelder's quality department. Establishing uniform reporting protocols allows stakeholders to easily interpret and compare quality assessment data, facilitating more informed decision-making processes. This standardization also enhances communication and collaboration within the organization, leading to greater efficiency and alignment of quality objectives.

Structured Evaluating Process
Establishing a structured process for evaluating and prioritizing suggested improvements is essential for driving continuous quality improvement initiatives at Van Gelder. By leveraging insights from the PQI cycle and data analysis, this process enables the identification of actionable improvement opportunities. Through systematic evaluation and prioritization, Van Gelder can focus resources on initiatives with the greatest potential for impact, ensuring a targeted and effective approach to quality enhancement.

Master Data Management
In addition to the aforementioned projects, implementing a master data management system is crucial for maintaining data integrity and consistency across Van Gelder's

quality department. By centralizing and standardizing data storage and management processes, this system ensures that accurate and up-to-date information is readily accessible to stakeholders. Furthermore, master data management facilitates data governance and compliance efforts, mitigating the risk of errors and inconsistencies in quality assessment processes.

Collectively, these projects represent strategic investments in Van Gelder's quality assurance capabilities needed to implement the complete PQi cycle to enable the organization to proactively manage quality risks, optimize operational processes, and deliver better projects for its customers. An overview of all the high-level action plans related to the PQi cycle can be seen in the figure below (fig. 4.6)

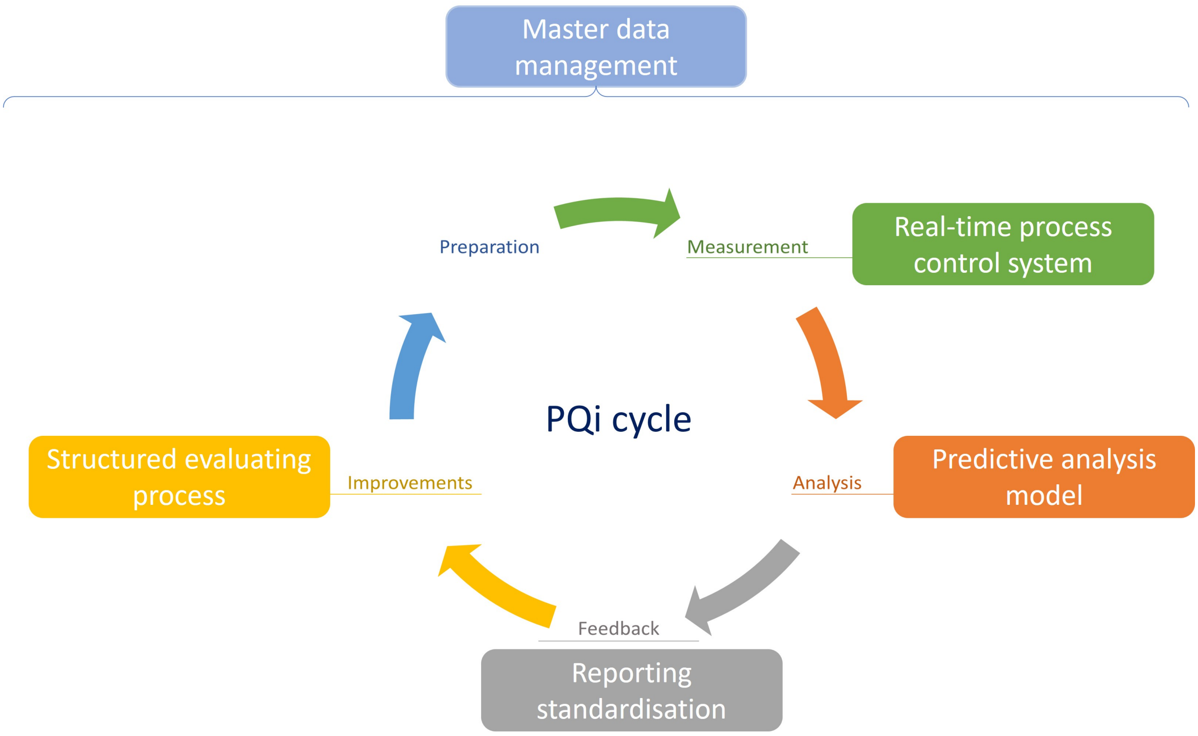


Figure 4.6: PQi cycle with high-level action plans

4.3.2 Architecture constraints and guardrails

This chapter describes the architecture constraints and guardrails that define the structure for the evolution of the EA4DT for the PQi initiative. The constraints and the guardrails give an understanding of the conditions under which the architecture (evolution6) is operated[51]. Architecture constraints define the limitations and boundaries within which the EA4DT must operate, while guardrails provide guidance and principles to ensure adherence to desired outcomes and standards.

4.3.2.1 Architecture Constraints

As stated above, the architecture constraints are the boundaries and limitations. These constraints could arise from different sources, such as technical, regulatory, and time-based factors [52]. Understanding and managing these constraints is essential for developing an effective EA4DT. The key constraints for this EA are listed below:

Technical Constraint
The first important constraint is that all changes made to current digital processes and future changes that influence the digital process have to go through the IT, Data, and Information departments of Van Gelder. For example, if the layout or content of the reports needs to change, this also holds for changes made to storing data or extracting data from the Van Gelder data hub. A second technical constraint is that the data generated by the system is secure and meets the requirements set by the IT department of Van Gelder. The final technical constraint is that the generated data can only be uploaded to the data hub once per day.

financial constraint
The second constraint is that the cost of making changes to the system needs to be financially beneficial. This also holds for indirect financial benefits, for example, if the systems are more accessible to interpret or if it increases the acceptance of their use

Business constraint
Another constraint is the system's usability. To keep the system usable and accepted, changes should be as minimal as possible. This means that people have to adjust their current working methods as little as possible. If more significant changes are needed to improve the system, this has to be adequately discussed with the system's end users, and training should be offered to keep the system understandable.
Another business constraint is that Van Gelder owns the system and its infrastructure. Companies may be called in to implement changes, but control must remain with Van Gelder.

4.3.2.2 Architecture Guardrails

Architecture guardrails are principles, guidelines, and best practices that provide direction and support for the design and implementation of the EA. These guardrails help ensure consistency, alignment with strategic objectives, and compliance with architectural standards. The following architectural guardrails are established for this project.

Alignment with Business Objectives
The EA4DT must be closely aligned with the organization's business objectives, priorities, and strategies. This guardrail ensures that architectural decisions contribute to the achievement of the (overall) business goals

Modularity and Flexibility
The EA4DT should be designed with modularity and flexibility in mind to accommodate future changes, updates, and expansions. This guardrail promotes agility, scalability, and resilience in the face of evolving business requirements and technological advancements.

Security and Privacy by Design
Security and privacy considerations should be integrated into the design and implementation of the EA4DT from the outset. This guardrail should adhere to Van Gelder's IT department's current safety standards/protocols. This protects sensitive data, systems, and assets against cyber attacks and unauthorized access.

Standardization and Consistency
The EA4DT should adhere to standardized architectural patterns, frameworks, and conventions to ensure consistency, interoperability, and ease of maintenance. This guardrail facilitates collaboration, reuse, and interoperability across different components and systems.

Continuous Monitoring and Optimization
The EA4DT should be subject to continuous monitoring, evaluation, and optimization processes to evaluate its ongoing relevance, performance, and effectiveness. This guardrail will help to achieve a culture of continuous improvement and adaptation to the changing business needs of Van Gelder and market conditions.

4.4 Architecture Outline

In Section 4.3, four high-level projects were highlighted to implement the PQi cycle. Due to the time limit of this research and the initial project from the problem description, the first project, a Real-time Process Control System, is chosen. Besides the fact that this was the initial project, it also serves as the basis for the other projects defined in the high-level action plan. Based on the results of this project, the different projects can be initialized simultaneously or one by one in a later stage.

4.4.1 Project requirements

As stated in the first research objective, this report section aims to determine the project requirements for the real-time process control system. To do so, we first need to have a better understanding of the research question *How to establish an Enterprise Architecture framework within a road construction company to accommodate a real-time process control system that ensures consistent data collection across diverse projects (both inter and intra projects), facilitates flexibility for customization, and promotes future-proofing in the absence of ready-to-use solutions in the market?* Certain terms require further explanation to clarify what is meant in this specific case before the project requirements are drawn up. These terms are consistent data collection and flexibility for customization. These have already been mentioned in section 2.2 but have not been properly defined yet.

1. **Consistent data collection across diverse projects**
   The word consistent is defined as *Constantly adhering to the same principles of thought or action* [53]. Data collection will thus entail ensuring that the data is collected through consistent principles of action. For this research, it means that the data is collected using standardized methods and protocols across projects, ensuring uniformity and reliability in the collected data. The second part of the term, diverse projects, comprises two project relation types. These are inter and intra-project data collection. We defined inter-project data collection as data collection from the same project teams but from different projects. We defined intra-project data collection as the collection of data for multiple projects running simultaneously.

2. **flexibility for customization**
   Flexibility is defined as: *Capable of being bent, admitting of change in figure without breaking; yielding pressure, pliable, plaint* and customizable is defined as: *That can be changed or produced to suit particular requirements or preferences* [54, 55]. For this project, the customization holds that changes can be made to the system according to Van Gelder's requirements, and flexibility ensures that the system keeps functioning with these changes.

3. **Real-time Data Collection Requirement**
   To ensure efficiency, the process control system should be able to gather information in real-time from all the individual nodes [1-10 seconds].

4. **Data Storage and Retrieval Requirement**
   The system should have enough backup capacity to retrieve data quickly and safely. It can be used by authorized personnel on analytical and decision grounds when needed.

5. **Data Integration Requirement**
   The system should integrate seamlessly with existing data management systems and databases within Van Gelder's infrastructure to ensure data consistency and coherence across different platforms.

6. **User Access and Interface Requirement**
   The system should provide a user-friendly interface for the end users and should provide the right information depending on the end user.

7. **Alerting and Notification Requirement**
   The system should have capabilities to generate alerts and notifications in (near) real-time based on predefined thresholds or anomalies detected in the collected data to enable quick responses and possible needed interventions.

8. **Data Security and Compliance Requirement**
   The system should adhere to official data security and compliance standards. Thus, the data should only be accessible to people with the proper clearance.

9. **Scalability and Performance Requirement**
   The system should be scalable to accommodate future expansion or increased data volume while maintaining functionality.

10. **Reliability and Redundancy Requirement**
    The system should be designed with redundancy measures and fail-over mechanisms to ensure uninterrupted data collection and feedback.

11. **Resolution requirement**
    Each cell within the real-time process control system should have a minimum resolution of 25 by 25 centimeters. This ensures precision and accuracy in data representation, allowing for detailed monitoring and analysis of the construction process at a granular level.

12. **Data quality requirement**
    The system must ensure high data quality throughout the data collection process. This includes measures to minimize errors and inaccuracies in the collected data. Quality control mechanisms should be in place to validate and verify data integrity on-site to ensure that the information used for analysis and decision-making is reliable and trustworthy. Additionally, the system should provide data cleansing tools to maintain the data repository's overall accuracy and reliability. The error margin for the resulting data can not be more than 10%.

4.4.2 selected project scope

The selected project scope has been mentioned shortly in section 3.4.3. The high-level action plan of developing a process control system was chosen to complete the remainder of the EA4DT framework to keep the research feasible. One of the reasons it was chosen is that it was in line with the company's overall objective and could serve as the basis for the rest of the high-level action plans. The process control system is limited by the actual paving process itself. It does not include milling, asphalt cleaning, applying the adhesive layer, asphalt transport, or asphalt maintenance. These individual stages of the asphalt process could be added later in the project to give a complete overview. Still, the process control system includes only the paving and compacting stages to keep this project feasible. The process control system will be worked out as a conceptual model; no physical system will be developed or tested within this research. The conceptual model will be validated through the decision-centric architecture validation method. It will cover all the requirements in detail except for the data security and compliance requirements.

5
Conceptual architecture

The conceptual architecture is the next phase for the EA4DT. The conceptual architecture is based on service-oriented design and implementation. Based on the architecture context, - vision, and - outline, the goal is to define the business services, portray the business service collaboration, derive the required information services, and derive the necessary platform services [39]. All this will form the application layer in the architecture and connect the business layer with the technology layer. The architecture outline listed all the project requirements. The requirements, however, are functional; for the conceptual architecture, they must be translated into services. Based on the functional requirements, the following four project requirement services are defined: real-time notification service, analysis service, reporting service, and data storage and management service. All services are described below with the corresponding characteristics. Figure 5.4 shows the new target architecture with the added services. The business and application levels are merged into one new target architecture.

5.1 Real-time information service

This service must ensure real-time notifications and visualizations are sent to relevant stakeholders (such as project managers, on-site executors, operators, or on-site personnel). The notifications enable them to respond to critical events or anomalies detected during construction. These notifications should be customizable based on predefined thresholds and user preferences, ensuring timely response to potential issues to allow responding. The visualizations, like the TCP and the CPM, must always be visible for the machine operators and on-site executors. Table 5.1 shows the characteristics of the real-time information service, whereas figure 5.1 shows the corresponding service architecture.

Table 5.1: Real-time Notifications Service

Characteristic	**Description**
Objective	Enable real-time information for stakeholders based on predefined thresholds or anomalies detected in the collected sensor data, facilitating quick responses and interventions and continuous information in the form of visualizations.
Business Process	The process of monitoring and visualizing sensor data in real-time, detecting anomalies, and triggering notifications to relevant stakeholders.
Business Roles	Stakeholders involved in the real-time monitoring and decision-making process, such as the on-site executor and operators.
Business Objects	Real-time notifications and visualizations, like the TCP and the CPM.
Information Systems Service	Real-time Data Monitoring and Alerting System.
Technology Services	Notification Engine and Visualization Engine with Messaging Platforms for information delivery.

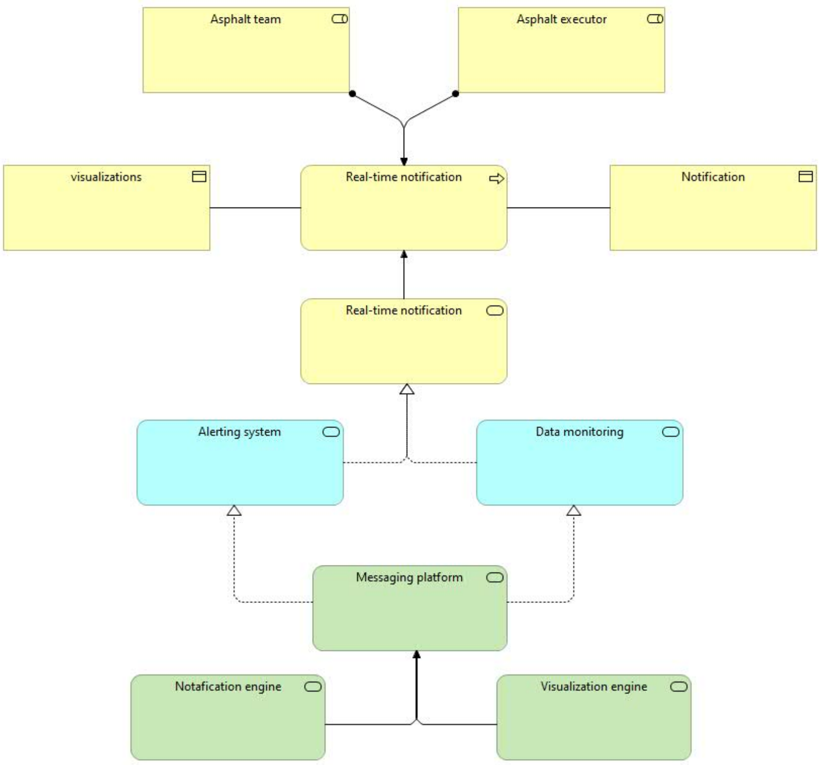


Figure 5.1: real-time notification service architecture

5.2 Analysis service

Implement an analysis service that performs data analytics on the collected project data to enable process learning and identify trends within the construction projects. This service should include functionalities such as statistical analysis, trend detection, and anomaly detection to optimize project performance. Table 5.2 shows the characteristics of the analysis service. Since the analysis and report service are highly correlated, both architectures are merged and are shown in figure 5.2.

Table 5.2: Analysis Service

Characteristic	**Description**
Objective	Provide analytics capabilities to derive insights from collected project data, supporting trend detection and process optimization possibilities.
Business Process	The process of conducting analytics on project data, including data pre-processing, modeling, and interpretation of results.
Business Roles	Data analyst responsible for conducting data analysis and interpreting results.
Business Objects	Analytical reports with insights and recommendations derived from project data analysis.
Information Systems Service	Project Analysis.
Technology Services	Predictive Analytics Processing Engine, Data Enrichment Services, Analytics Development Environment Access.

5.3 Reporting access service

Create a reporting access service that allows access to customized reports and dashboards summarizing key performance indicators (KPIs), trends, and actionable insights derived from the collected data, depending on the stakeholder. These reports should be accessible to various stakeholders, providing transparency and facilitating data-driven decision-making at different levels of the organization. Table 5.3 shows the characteristics of the reporting service, whereas figure 5.2 shows the corresponding service architecture.

Table 5.3: Reporting Service

Characteristic	**Description**
Objective	Enable stakeholders to access predefined and customizable reports based on processed project data, promoting data-driven decision-making and performance monitoring.
Business Process	Access itself has no process but allows for operational and strategic planning and access to project delivery control reports for clients.
Business Roles	The process manager, operational manager, ATs, asphalt executor, and the head quality department collaborate on generating standard reports and using the information for process improvement
Business Objects	Standard and custom reports with data visualizations and dashboards for internal use.
Information Systems Service	Reports access provides the application user interface to access analytics reports.
Technology Services	Reports Web Portal Access, Integration with Data Visualization Tools, Self-service Data Access.

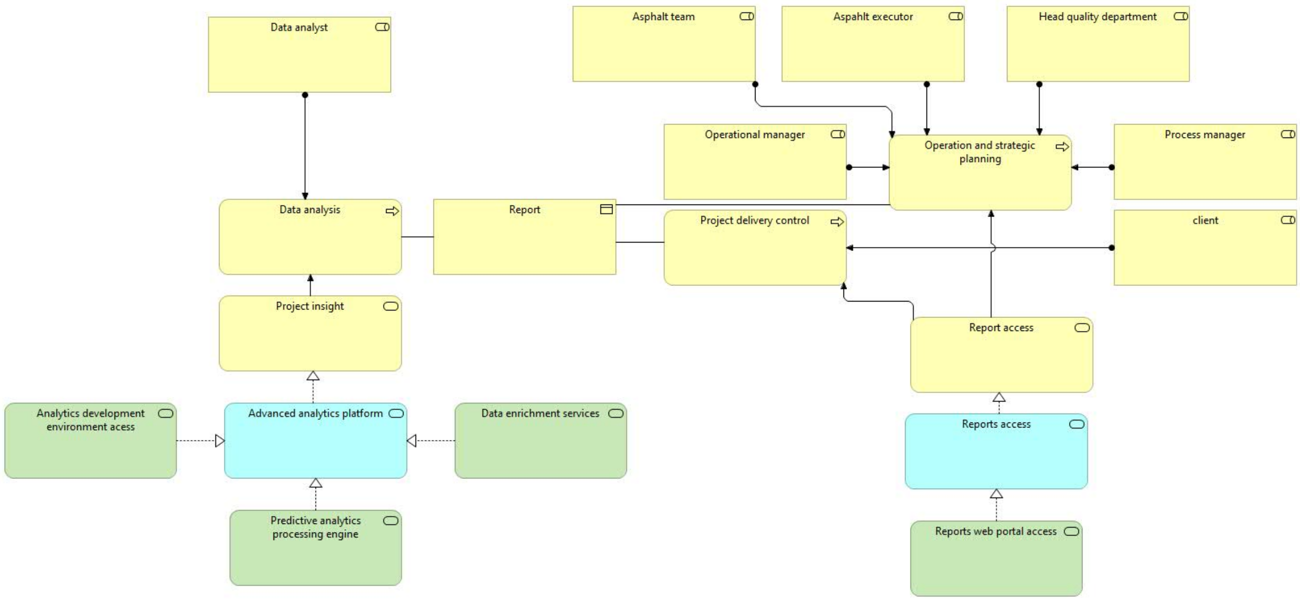


Figure 5.2: analysis and reporting service architecture

5.4 Data storage and management service

Develop a robust and reliable data storage and management service that securely stores and organizes the collected project data in the data hub of Van Gelder while adhering to the security guardrail set in section 4.3.2.2. This service should include data ingestion, storage optimization, indexing, and data lifecycle management to ensure efficient handling of large volumes of data generated by construction projects. Data governance policies and access controls must also be implemented to maintain data integrity. Furthermore, the data quality and format should be aligned with relevant regulations. The service should also support seamless integration with existing data management systems and databases within Van Gelder (as stated in section 4.3.2.2), enabling data consistency and coherence across different platforms. By implementing this service, Van Gelder can establish a reliable and scalable foundation for storing, accessing, and managing project process data, facilitating informed decision-making, and enhancing operational efficiency within the organization. Table 5.4 shows the data management and storage service characteristics, whereas figure 5.3 shows the corresponding service architecture.

Table 5.4: Data Storage and Management Service

Characteristic	Description
Objective	Develop a robust data storage and management service that securely stores and organizes the collected project data in the data hub.
Business Process	The processes of data ingestion, storage optimization, indexing, and data lifecycle management to efficiently handle large volumes of construction project data.
Business Roles	IT, the process manager, and Data Information (DI) department are responsible for overseeing data storage, organization, and access within the organization.
Business Objects	Data storage structures and data lifecycle management policies.
Information Systems Service	Data Storage and Management Platform.
Technology Services	Data Ingestion Engine and Data Lifecycle Management Tools.

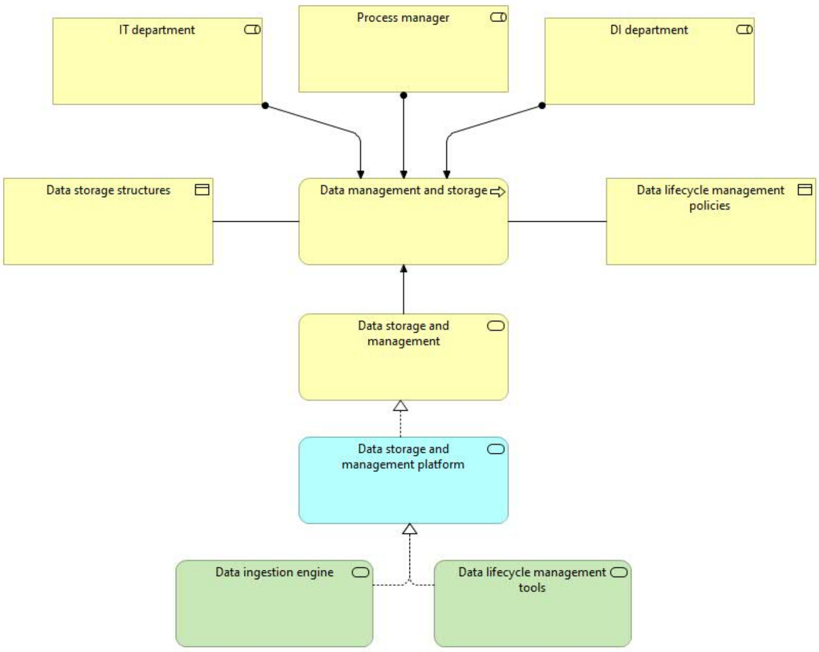


Figure 5.3: Data storage and management service architecture

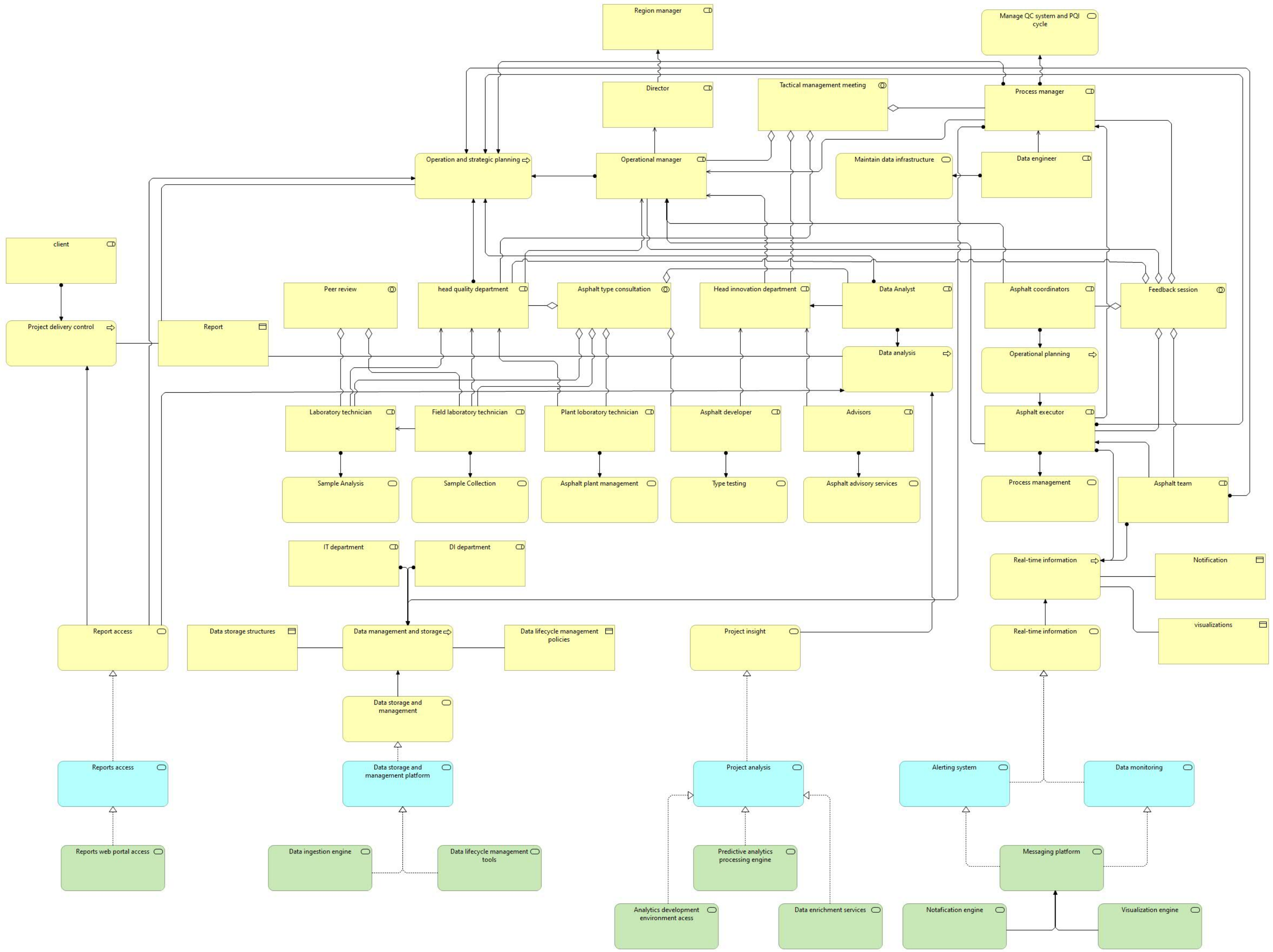


Figure 5.4: target architecture with business and application level

6
Logical architecture

For the logical architecture in this research, we will examine the data needed for the process control system based on the requirements and literature. From this, a data ontology will be developed that can be used for the remainder of this project. Finally, a short section will be dedicated to data management. All this will serve as a base for the logical architecture.

6.1 Data ontology

This research section will cover the data needed for an effective process control system and the corresponding data ontology. A data ontology can be defined as: *"A description of data structure–of classes, properties, and relationships in a knowledge domain. It is meant to serve as a basis for instances of knowledge graphs, ensuring data consistency and understanding of the data model."* [56]. Thus, a data ontology serves as a means to properly structure the data and define the right relations between them to get a structured overview for better understanding and implementation.
As mentioned in chapter 1, Makarov has developed a proof of concept for a process control system, provided a detailed overview of the data needed, and included a data ontology. Recently, Sadeghian provided a detailed data ontology on pavement life-cycle management [2]. In his Engineering Doctorate thesis, Sadeghian developed an ontological model for pavement projects, including the design, construction, operation, and maintenance. Although the ontological model is conceptual and did not undergo practical testing, it has a high level of detail and is validated by stakeholders. The ontology has been developed with the input of stakeholders from contractors, clients, and third parties based on semi-structured interviews and existing literature.
both ontological models in the papers from Makarov and Sadeghian are seen as supplementary to each other and thus will be used together to develop a data ontology for the process control system. As mentioned earlier, Sadeghian developed an ontology for the entire life cycle of the pavement. In this research, however, only the construction section of his ontological model will be considered.

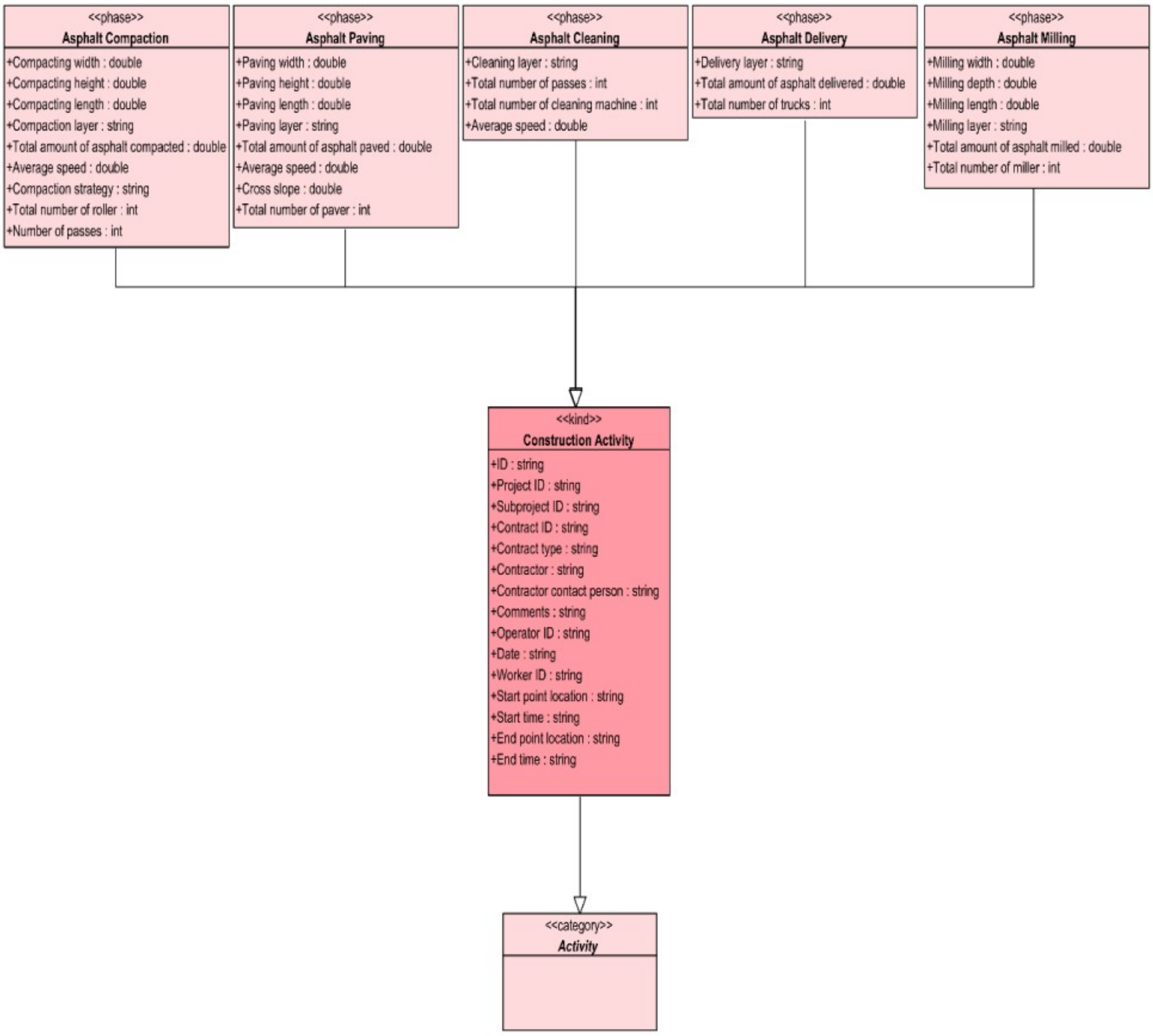


Figure 6.1: Ontological model Sadeghian [2]

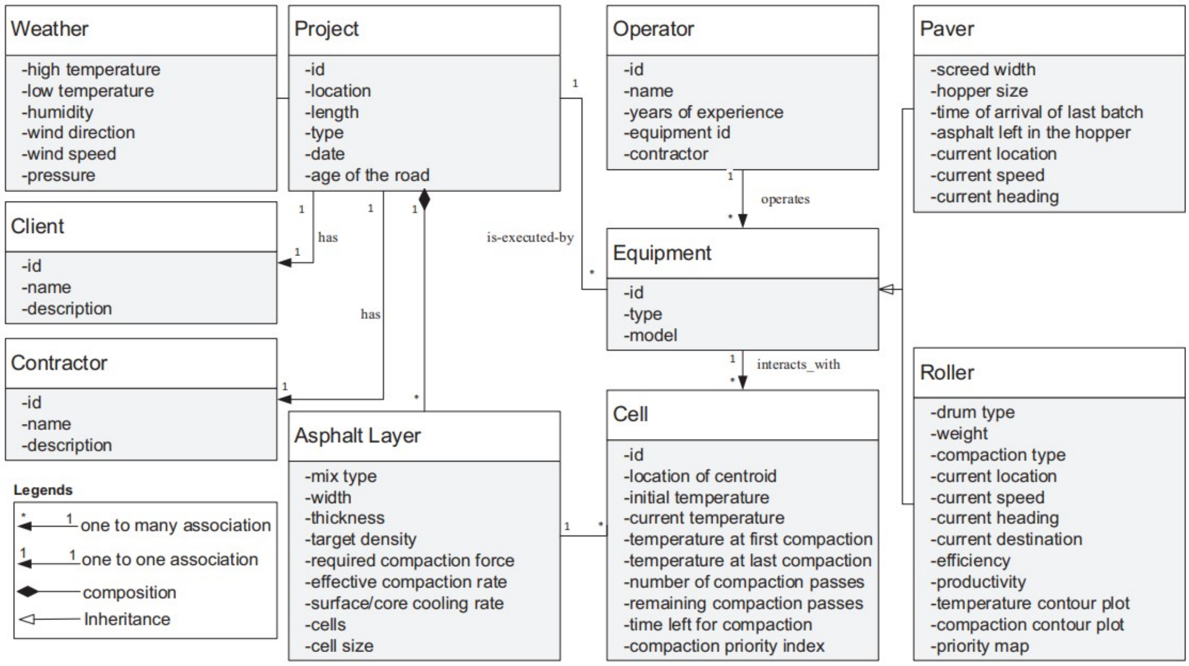


Figure 6.2: Ontological model Makarov [3]

The next step is to analyze and compare both ontological models to create a new, more complete ontological model. It is important to keep the development purpose of each ontological model in mind. Makarov's ontological model was developed for the compaction guidance system, whereas Sadeghian developed his ontological model based on the input of contractors, clients, and third parties. Therefore, the ontological structure of Makarov's model will be used since this is on a system level, and this model will be supplemented by Sadeghian's ontological model. As mentioned in section 4.4.2, The selected project does not include milling, asphalt cleaning, applying the adhesive layer, asphalt transport, or asphalt maintenance.
Figure 6.3 shows the revised ontological model. Before further explaining the revised model, it has to be said that some types of given attributes are not entirely accurate. This concerns the missing type 'string' in the program Enterprise Architect. All the 'string' types are changed to 'char'. The initial attributes of Makarovs model are left in the revised model since they were either also mentioned by the ontological model of Sadeghian or needed to make the process control system functional. However, attributes are added based on Sadeghian's ontological model. In the weather class, the weather station's location, altitude, precipitation, and comments are added. The first three are additional information that can be used for better analysis. The latter gives the opportunity to add comments, for example, if the station fell over or had to be moved during the paving process. This can help filter out unwanted data. In the project class, project id, sub-project id, contract id, contract type, number of pavers, number of rollers, comments, start time, and end time. Here, Sadeghian has added and proposed many attributes. The first six attributes help to define the project details better. In contrast, the comment may help to describe some phenomenon, like an accident or a machine that

broke down, which can allow for better data analysis since these things might be hard to or impossible to notice/determine when someone looks at the data only. The start and end times were added since they could be used for statistical analysis. In both the client and contractor classes, comments and a contact person attribute are added. The comment attribute might seem redundant since there is already a description attribute for both. In contrast, the description might be general and lengthy; the comment can be specific and relatively short. The contact person attribute is added in case more information is needed or a change needs to be reported to the corresponding party. In the paver class, only the total amount of asphalt paved is added as an attribute. This is very useful since it can help the paver operator or the executor see the progress and determine if the project needs as much or more asphalt than anticipated. The executor could get an alert if the amount of asphalt over the expected amount exceeds a certain amount. In such a case, the executor can contact the asphalt plant to bring more or less asphalt. In the roller class, no new attributes are added. Sadeghian had average speed, compaction strategy, and total amount of asphalt compacted as attributes in his asphalt compaction phase class, which were not in the model of Makarov. For this revised model, they are left out since they do not contribute to a better functioning or more complete process control system. In the operator, equipment, and cell classes, no new attributes are added. Finally, cross slope, length, and layer attributes are added to the asphalt layer class. Sadeghian also had the number of lanes as an attribute, but this does not seem to contribute to the use case of this ontological model, so this is left out.

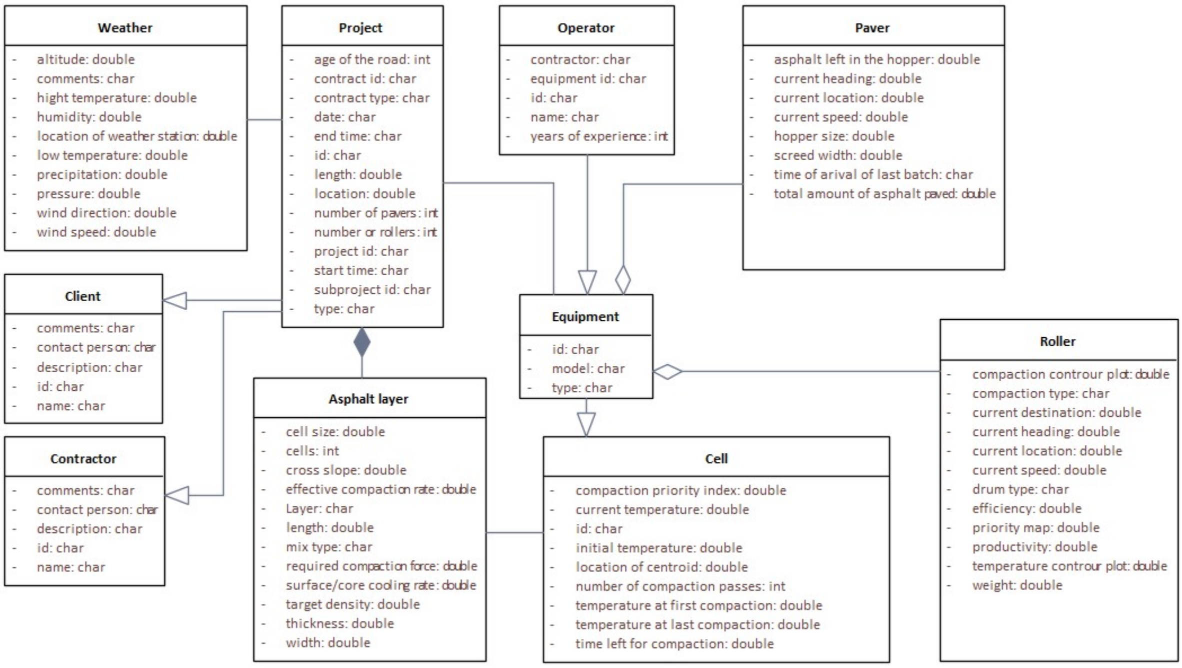


Figure 6.3: Revised ontological model

6.1.1 Sensor nodes

The revised ontological model presented in figure 6.3 shows all the classes and attributes within these classes. In order to make a functional process control system within a functioning PQi cycle, all these attributes need to be collected and stored for real-time use or post-project analysis. In this section, all sources/sensors will be mentioned in order to collect the attributes. First, a distinction needs to be made between attribute data that is collected on the construction site and attribute data that can be captured in Van Gelder's portals or database. Continuous data that needs to be collected on the construction site will be called dynamic data. Data that is predetermined and not likely to change during construction will be called static data. Client data, for example, will not change during the project and can be imported from Van Gelder's data hub. The two sections below will cover all the classes in the revised ontological model (fig. 6.3) by categorizing them as static or dynamic data and determining the data source, if possible. It is important to note that a class categorization does not mean that all attributes in that class belong to that category. Furthermore, it is essential to add that no matter the type of data group, It is possible to change attributes. Staff with the proper access are able to change the contract type or the operator on the machine after completing the project if they spot a mistake.

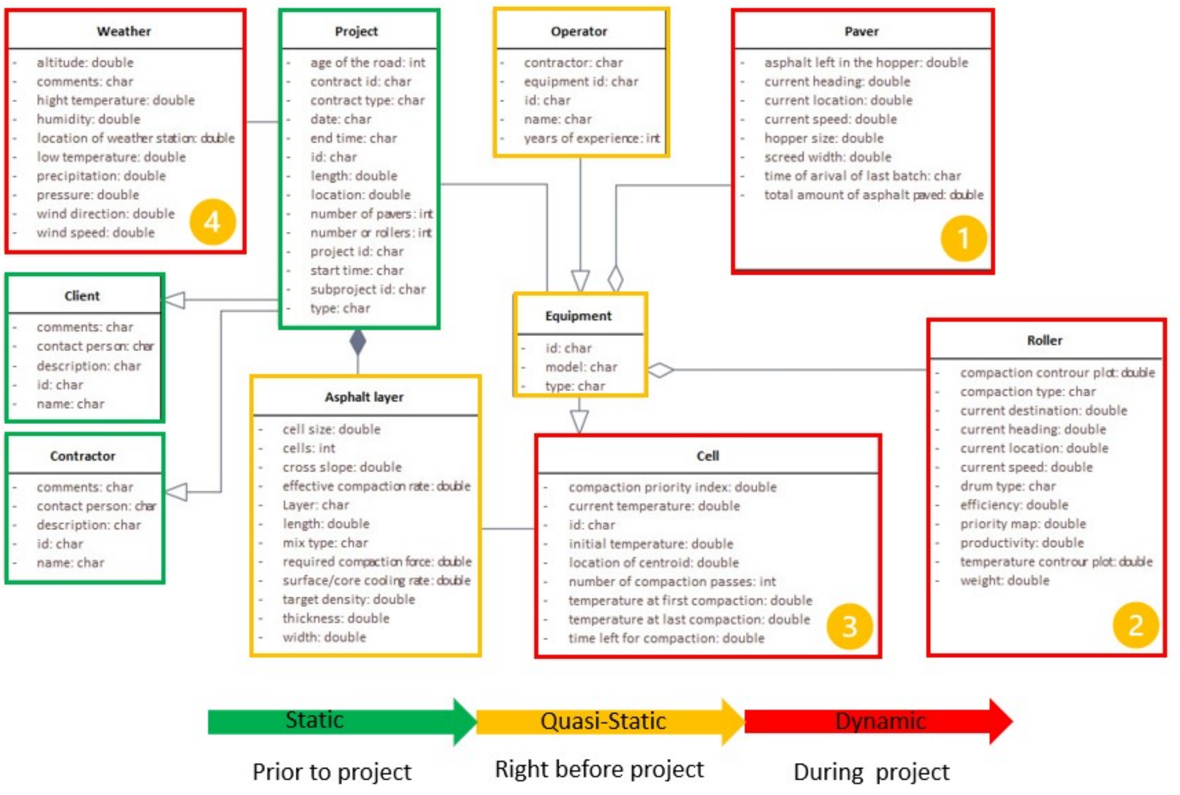


Figure 6.4: Revised ontological model with data type indication

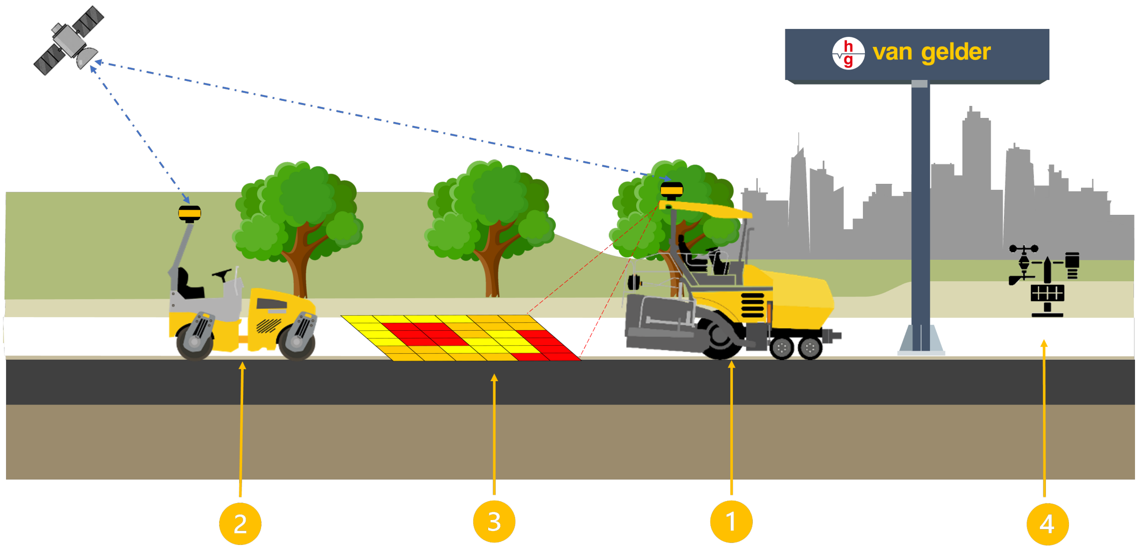


Figure 6.5: Dynamic sensor data overview

6.1.1.1 Static data

As mentioned in the section above, static data is data that is predetermined and not likely to change. The static classes are highlighted green in figure 6.4. These are the project -, client -, and contractor classes. The project class has some attributes that are dependent on the construction process and can thus only be collected after the completion of the project. Start - and end time and length have to be added after the completion of the project. All the attributes from the static data classes, except for the construction-dependent attributes, can be collected from Metacom. This is the portal that Van Gelder uses, which contains all the project information.

Quasi-static data

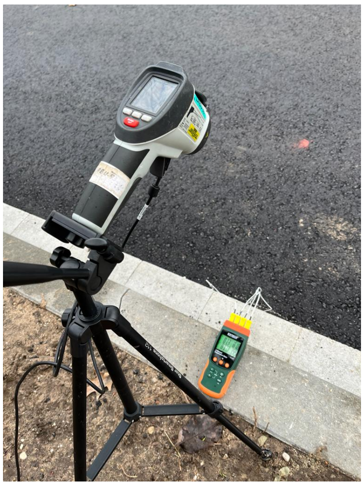
Within the static data group, we will also introduce a quasi-static data group. Like the static group, the quasi-static group contains predetermined data. The difference is the change in data. The quasi-static data class contains the asphalt layer -, operator -, and equipment classes. These classes are highlighted yellow in figure 6.4. Like the static data group, these classes are planned and thus static, but in reality, these classes are known to change last minute. It could, for example, be that on the day of construction, different equipment is used due to unscheduled maintenance or a swap of machinery among two projects. So, like the static group, the attributes from the quasi-static operator and equipment classes are collected from Metacom. The characteristics from the quasi-static asphalt layer class are collected from a portal called 'PIM'. The cooling rate attribute in the asphalt layer class is dynamic and has to be added after the completion of the project. The system should, however, allow changes to be made to the system right before the construction process starts. This means that either the information in either portal should be changed or an authorized person is able to overwrite the data coming from the portals.

6.1.1.2 Dynamic data

The dynamic data classes are highlighted in red in figure 6.4. The dynamic data group is continuous data that is collected during construction. this group contains the weather -, paver -, cell -, and roller class. The weight attribute of the roller and the hopper size attribute of the paver are predetermined. Figure 6.5 illustrates the individual nodes for dynamic data collection during construction. Figure 6.6 shows all the sensors needed to collect the remainder of the dynamic data class attributes. Figure 6.6a shows the GPS sensor needed for all the classes. Figure 6.6b shows a thermal camera for the paver class. The paper also needs a screed width sensor and a layer thickness sensor. Figure 6.6c shows the temperature station to determine the cooling rate of the asphalt to determine the time left for compaction for the cell class. Finally, figure 6.6d shows the weather station, which registers all the attributes from the weather class.

(a) GPS sensor

(b) Thermal camera



(c) Temperature sensors

(d) Weather station

Figure 6.6: Dynamic data sensors

6.2 Data processing and data storage

Above, we have presented and elucidated the data ontology. Now that we have identified the necessary information to make a functional process control system, it is imperative to establish the data infrastructure for data processing and data storage. This covers the processing and storage location of the data. The processing of data covers the analysis of the data and the generation of relevant visualizations, like the TCP, CCP, and the CPM, see figure 1.4. Makarov included the data network as part of the processing, but in this paper, the data network will be explored later in this paper [7]. Data storage is responsible for string the entire data set of the entire duration of the project for post-project analysis. This section of the report will explore local, central, and hybrid options for data processing and storage. One might say that it would be more logical to work out the data acquisition and the data network before looking into the data processing and storage. However, the data processing and storage are explored first, since this is a more important part of adhering to the set requirement and the data acquisition and network should be based on this.

6.2.1 Data processing

The data processing is responsible for transforming the data from the individual nodes from each project to the desired outcome, such that the roller operator is able to see a real-time CPM and the paver operator is able to see the TCP, **more based on the ontology**. The processing of the data from the individual nodes on the project site has to be done by a computer and can either be done locally on the project site or centralized. Even though the cloud-first strategy was mentioned as a technology principle in section 4.2.1, more in-depth analysis is needed to select the best approach for data processing of the system.

Local processing
Local processing would require a local processing unit to analyze the data and generate relevant visualizations, see 6.7 For van Gelder, there would be the need for five local processing units, which from now on will be called servers. The processing power needed for this process control system could fulfilled with a mid-range laptop [7, 9]. Local processing is not in conflict with any architectural constraints set in section 4.3.2.1; this implies that for the technical constraint, it is assumed that the system is able to adhere to the requirement of data security set by Van Gelder.
In order to see if the use of local processing would be suitable for the process control system, all the requirements set in section 4.4.1 will be checked and further explained if needed. An overview of all the requirements and if it meets the requirements can be seen in table 6.1.
Having a separate server for each project does allow for consistent data collection across diverse projects. Interfering with one project would never affect another project at the same time. It, furthermore, gives a great deal of flexibility for customization to the

system since changes could be made fast and without interfering with other projects running at the same time [57, 58]. Another benefit of a local server is the low latency. With the server close to the project, data can be sent back quickly. Data storage and retrieval requirements will be discussed in section 6.2.2. Requirement numbers 5-7 are related to data coherence, user interface, and alerting and notification, which are all dependent on the algorithm and its design choices. It is assumed that both server types are able to perform this algorithm and are thus left out for this comparison. As stated before, the data security and compliance requirement was set as one of the guardrails and mentioned that it is outside this project's scope. The scalability and performance requirements could pose a risk for local servers. Whilst the latency might be low due to its close proximity, the processing might eventually limit the real-time analysis with future add-ons. For example, if more sensors are added to the system, more parts of the process, like transportation, are added to the system, or if there is a need for higher data quality, the server might have to be upgraded. This is possible but can become costly since it has to happen for all the ATs [57, 58]. It is not a limiting factor but might become financially limiting if the company needs to adhere to the financial guardrail (section 4.3.2.2). The same holds for the reliability and redundancy requirements. If a local server fails, it would require a backup, and only a local backup could keep the system functioning. A local server can be as reliable as a centralized server but requires maintenance. The cost to maintain five local servers (plus additional backup servers) would outweigh the cost to support one central server [57, 58].

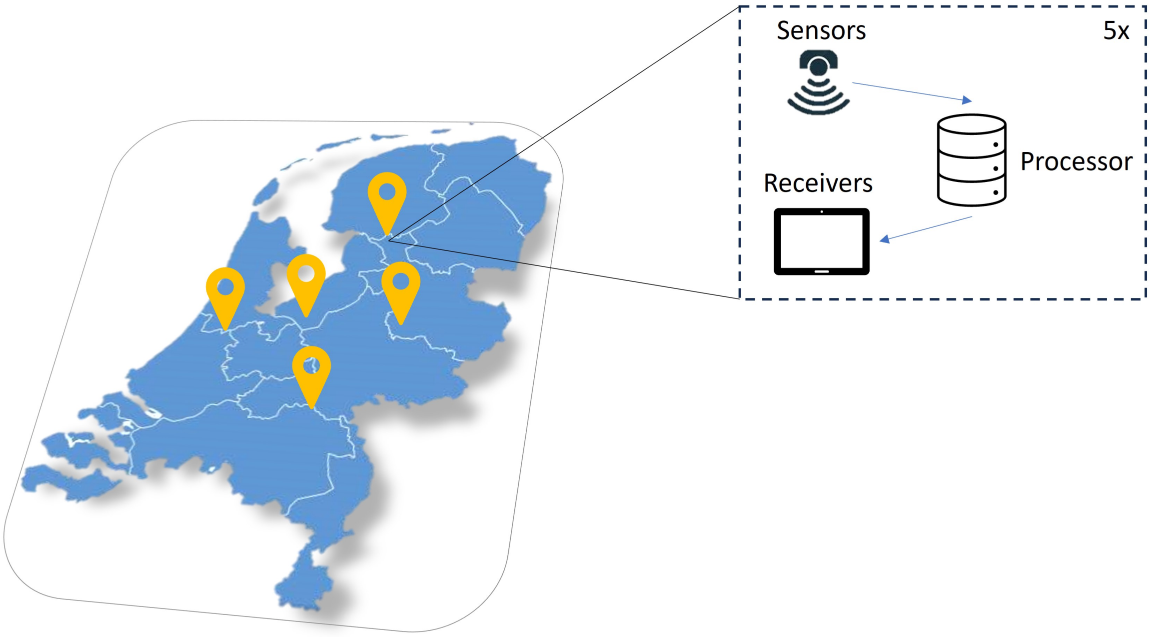


Figure 6.7: Local processing

Central processing
For central processing, the processing power is not present at each project location, but the server is located in one place, and all the data is sent to this server. A benefit of having a central server is cloud computing. With cloud computing, the server is owned by a third party and can be a shared asset [57, 59]. The local server will be referred to as cloud computing from now on. Cloud computing has, like a local server, its positives and negatives. In order to see if the use of cloud computing would be suitable for the process control system, the exact requirements, as in the section above, will be checked and further explained if needed. The results of this can be seen in table 6.1. Cloud computing gives the ability for consistent data collection across diverse projects, but a change in the protocol might affect multiple rather than one project in the case of local servers. The degree of flexibility for customization with cloud computing is also lower compared to having local servers; since the central server is owned by a third party, changes made to the system will probably take longer to implement and will most likely be more expensive [57, 58]. This relates more to the customization part of the requirement; the level of flexibility would be the same for cloud computing. As stated before, the data security and compliance requirement is left out of this project's scope. Still, it is worth mentioning that the data security of cloud computing is, in most cases, handled by the operator of the server and, thus, not the responsibility of the IT department of Van Gelder [57, 58]. A huge benefit of cloud computing is the degree of scalability and performance. With cloud computing, there is no need for extra hardware in order to increase the processing power, and there is no need to invest in new servers if more ATs are being established in the organization [58]. If the process control system is expanded in the future, with the examples given in the paragraph above, the processing power can be increased to the desired level with ease. However, it is also the other way around; if less processing power is needed, it can be down-scaled to lower the operating cost. Another huge benefit of cloud computing over a local server is the reliability and redundancy. Cloud computing is more reliable than local servers since they can ensure a certain level of uptime with a so-called 'cloud service-level agreement (SLA)' [58]. Such a SLA ensures that a certain level of service is maintained, guaranteeing a certain level of reliability, responsiveness, and availability [60].

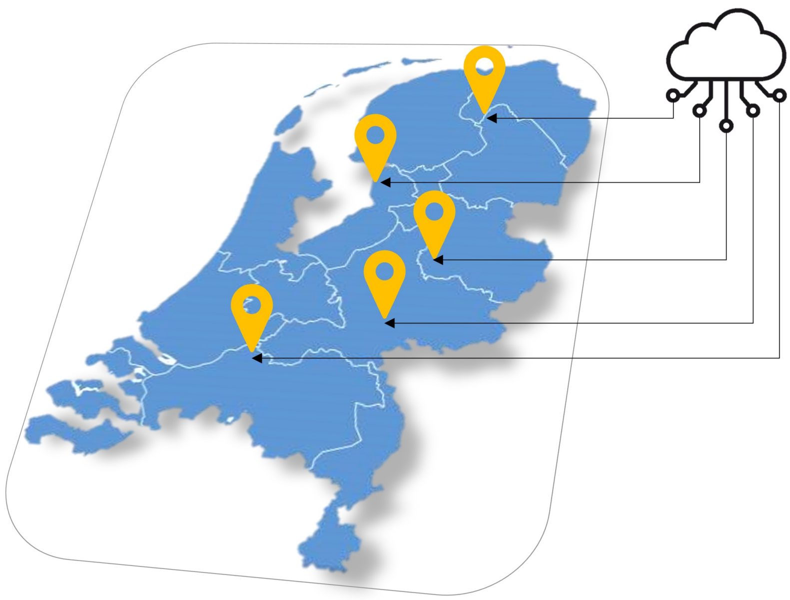


Figure 6.8: Central Server - Cloud computing

Both a local - and a cloud server could be used to develop a functional process control system. Both server options can meet the functional requirements set in section 4.4.1. It comes down to the investment and operating cost to generate the desired performance of the system, which makes the difference. The local server has the advantage over cloud computing when it comes to flexibility. The local server, however, loses out over cloud computing when it comes to security, scalability, performance, reliability, and redundancy. Thus, if both server options are compared, it has to be said that the cloud server is the most favorable option. For this reason, the remainder of this framework will be based on cloud computing for processing. This is also in line with the technology principle set in the architecture vision (4.2.1).

Table 6.1: Requirements processing

nr.	Requirement	Comparisson
1	Consistent data collection across diverse projects	Local = Central
2	flexibility for customization	Local > Central
3	Real-time Data Collection Requirement	Local = Central
4	Data Storage and Retrieval Requirement	Local = Central
5	Data Integration Requirement	Local = Central
6	User Access and Interface Requirement	Local = Central
7	Alerting and Notification Requirement	Local = Central
8	Data Security and Compliance Requirement	Local < Central
9	Scalability and Performance Requirement	Local < Central
10	Reliability and Redundancy Requirement	Local < Central
11	Resolution Requirement	Local = Central
12	Data Quality Requirement	Local = Central

6.2.2 Data storage

Besides the real-time functionality of the process control system, the data also needs to be available to the data analyst. An important constraint that has to be taken into account is the final technical constraint mentioned in section 4.3.2.1. This technical constraint holds that the data is only uploaded to the data hub of Van Gelder once per 24 hours. For van Gelder, this is done from 11 pm to 6 am. This means that the data is only available for analysis the next day. In this case, we assume that the data analyst is only authorized to work with data available to him from the data hub of Van Gelder. This is also in line with the security and standardization guardrails set in section 4.3.2.2. Since the data cannot be continuously uploaded to the data hub of Van Gelder, there needs to be another storage option to store the data for one day before uploading it at once to the data hub. The section above concluded that central processing would be the best option for the process control system. This gives two possibilities: either the data is stored centrally or locally. A third option would be hybrid data storage, where the data is both stored locally and centrally.
The hybrid data storage option would be the best option with regard to the requirements (sec. 4.4.1). The central data storage can be coupled to the same central processor used for cloud computing. This data storage option is favorable since an API can be used. An API is a mechanism that allows two software components to communicate with each other [61]. In our case, there is a connection between the central data lake and Van Gelder's data hub. All data is already in one place; thus, there is only a need for one such API. It furthermore would reduce the security risks. The advantage of additional local storage in combination with central storage is the redundancy. In the case of data loss in the central database, data can be retrieved from the local data storage. Figure 6.9 illustrates the central processing and hybrid data storage.

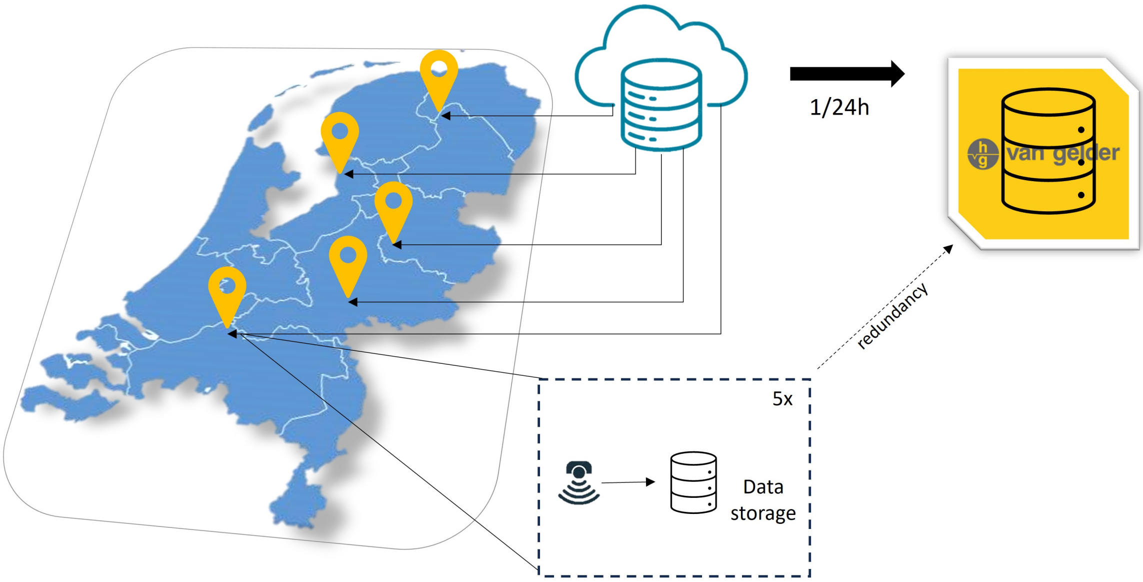


Figure 6.9: Central processing and hybrid data storage

6.2.3 Data network

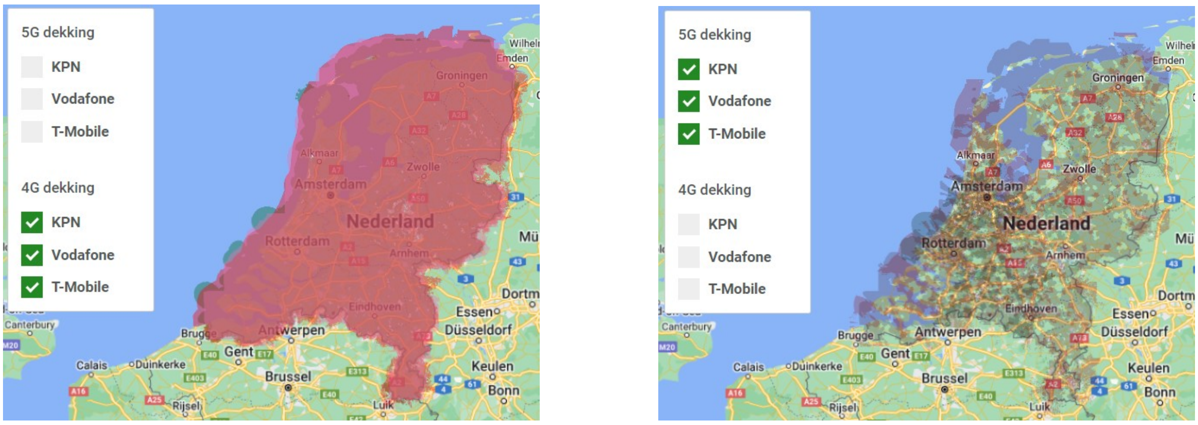
The previous sections defined the data that is needed with an ontological model and the method on how to process the data. To complete the real-time process control system, a data network has to be chosen in order to get the data from each project location to the cloud computing server. While options like Bluetooth (Personal Area Network (PAN)) and Wi-Fi (Wireless Local Area Network (WLAN)) may be suitable for local and stable networks, they would fall short in scenarios requiring broader coverage and reliability. With the use of cloud computing, the system architecture for the process control system turns out to be based on the principles of the IoT. IoT refers to the interconnected network of physical devices, sensors, and software applications that communicate and exchange data over the internet [62]. By using the power of IoT, it is possible to collect and analyze real-time data from various sources, enabling real-time feedback.

Table 6.2: Network Options for IoT

Network Option	Coverage	Data Speed [Mb/s]
2G	good	0.3 [63]
3G	good	42 [63]
4G	good	300 [63]
5G	limited	10^3 [63]
NB-IoT	good	0.25 [63]
LTE-M	good	1 [63]
LoRaWAN	good	0.05 [63]

In terms of network options, there are two main alternatives: Long Range Wide Area

Network (LoRaWAN) and cellular networks. LoRaWAN offers long-range, low-power wireless communication suitable for IoT applications in remote areas. On the other hand, cellular networks provide broader coverage and higher data speeds, making them suitable for applications requiring real-time data transmission over larger areas. The choice between these options will depend on the project requirements, coverage area, and data transfer speeds. The related project requirement for this section is the real-time data collection requirement. This requirement is not only based on the speed of the network but also on the processing power. It is furthermore hard to determine the exact amount of data produced by each node. 6.2 shows all the network options. The data speed refers to the download speed; the upload speed is generally lower than the download speed [64]. LoRaWAN is specially developed for the exchange of small amounts of data with a deficient energy consumption [63, 65]. The energy consumption is no limitation for this project since the sensors and transmitters can either be connected to the machine or the battery can be charged or swapped easily. Due to the low data speed, LoRaWAN will not be used for this project. Narrow Band (NB)-IoT and Long Term Evolution (LTE)-M are both based on the principles of 4G [63]. Both have a deficient power consumption but also have a rather low data speed [63, 65]. The same holds for the 2G network. The 2G network will also lose support from the Dutch providers in the coming years [66]. For this reason, these network types will also not be used for this project. The data speed of 3G should have enough speed to collect data from each separate node. The 3G network is not supported anymore and thus not suitable for this project [65, 66]. 4G and 5G both have enough data speed, but the coverage of the 4G network is higher compared to that of 5G; see figure 6.10 [67]. The network that will be used for this project is 4G since it has a high data speed of up to 300 Mb/s and good coverage in the Netherlands (fig. 6.10a).



(a) 4G network coverage

(b) 5G network coverage

Figure 6.10: 4G and 5G network coverage

van gelder

6.2.4 Technology architecture

In section 4.2.5, a target architecture has been developed on the business level for the architecture vision. Later, the high-level action plans were defined in section 4.3, and a new target architecture was developed with the needed services on the application level (Chapter 5). In section 6.1, a revised ontological model was designed to establish a baseline for the information needed for a functional process control system. This section of the report aims to develop the data architecture on a technology level to complete the logical architecture and update the target architecture. Figure 6.11 shows the architecture with application and technology level. The project node on the bottom right represents the information gathered on the construction site. These contain the roller -, paver -, weather -, and cell class, which includes all the physical sensors presented in 6.6 for the dynamic data collection. Each class(node) has its own gateway to connect the sensors to the cloud for processing. For the real-time information service, the cloud servers serve as the notification engine, visualization engine, and messaging platform. The notifications, TCP, and CPM are sent to the notification application and user interface for visualizations. From the cloud server, the data is also sent to a data lake to store the data temporarily, as shown in figure 6.9. Metacom and PIM are connected to the data lake with an API to store the relevant static and quasi-static data as shown in the revised ontology (fig. 6.4). All the project data from one day flows to the data ingestion engine, which is realized by the cloud server and entered into the data hub of Van Gelder. The data management server realizes the data life-cycle management tools to support the data storage and management service. The authentication and authorization server serves the reports and web-portal access and the analytics development environment access services for the power Business Intelligence (BI) and modeling and processing application. Both applications are fed by the data hub. Figure 6.12 shows the complete target architecture with the business -, application -, and technology level.

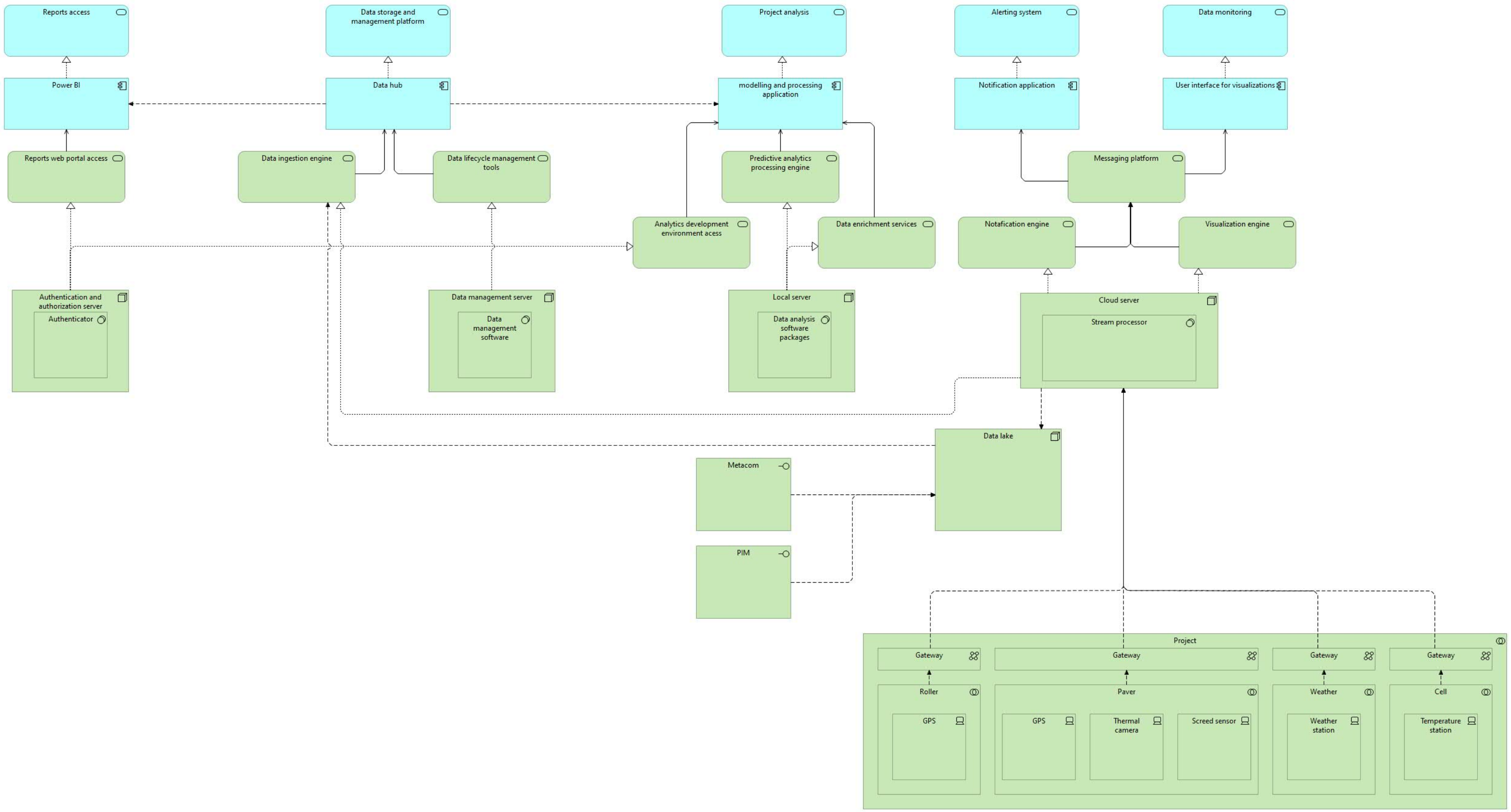


Figure 6.11: Target architecture with application and technology level

Figure 6.12: Target architecture with business, application, and technology level

6.3 Data Management

Introduction

Data management is a study on its own and has been mentioned already in the high-level action plan (sec. 4.3) as an individual project to complete the PQi-cycle implementation (4.6). Nevertheless, it is briefly discussed in this report because it also plays a vital role in the real-time process control system project.
Like Van Gelder, many organizations strive to empower the information of the different parts of their organization by sharing information and promoting collaboration. As the speed and efficiency of the underlying infrastructure improve, so does the organization's ability to control and take advantage of information assets through information sharing and collaboration. This collaborative environment not only enhances connectivity between systems but also fosters stronger working relationships among personnel, leading to more effective business management and a competitive edge. The multiplication of such data across different business applications within an enterprise has resulted in so-called "islands of information coherence" [68]. Where the enterprise is often composed of many applications referring to multiple, sometimes disparate sets of data that are intended to represent the same and vice versa. This highlights the need for a consistent view of data across business boundaries [68].
Master data is the most important thing for an organization. For the enterprise (asphalt organization) of Van Gelder, this is the data collected during construction and the (quasi) static data described in section 6.1 and represented in figure 6.4. These master data objects are the core business objects that are used in different applications across the organization, along with their associated metadata, attributes, definitions, roles, etc. MDM is a collection of best data management practices that orchestrate vital stakeholders, participants, and business clients. MDM integrates business applications, information management methodologies, and data management tools to establish the policies, procedures, and infrastructure necessary for capturing, integrating, and enabling the use of shared, precise, timely, consistent, and complete master data [68]. The use of MDM has several essential benefits: consistent reporting, improved competitiveness (faster integration of new data and system into the organization (flexibility for customization)), improved operational efficiency and reduced cost, Improved decision making, better spending analysis and planning, increased information quality, quicker results, improved business productivity [68]. Loshin describes that the implementation of MDM is challenging and must be seen as a program rather than a project. This section of the report will cover a general view of the MDM paradigms, architectures, and the management guidance

6.3.1 MDM paradigms and architectures

Loshin describes three MDM architecture styles: the registry, the transaction hub repository, and the centralized master repository [69]. There are three main factors that make up the implementation spectrum for the MDM architecture styles; see figure 6.13. Before continuing, it is essential to mention that the MDM architecture will be based on the data hub of Van Gelder and, thus, not on the real-time information service.

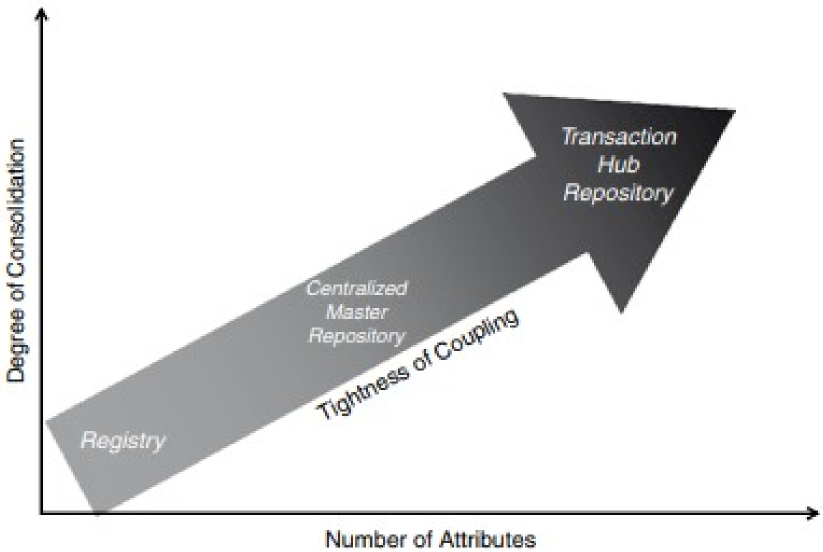


Figure 6.13: MDM architectural styles along a spectrum

The degree of consolidation is dependent on the amount of data sources and the need for a "single view" [69]. The number of data sources is low in this case. Based on the current target architecture (fig. 6.12), the data going to the data hub of Van Gelder is coming from 1 source (data lake). The need for a single view, however, is more important. This relates to the structure and uniformity of the data. It is essential that the data going into the data hub is the same for everyone who has access to the master data. The amount of attributes is relatively high and might become even higher if more attributes or classes are added in the future. The tightness of the coupling depends on the coupling between the applications and the master data repository. There is a tight coupling between the modeling and processing application and the master data in the data hub. The coupling between the power BI application and the data hub can be less tight but still should be tight enough so that changes made to the master repository can be seen in the reports generated by power BI. If all the individual requirements for the MDM architecture styles are combined, it can be said that the centralized master repository would be the best architecture for this use case. The architecture of the centralized master repository is shown in figure 6.14 below. The choice for this particular architecture is further supported by its characteristics, which are highlighted by Loshin. First of all, the harmonized Views of Master Objects: Van Gelder aims to establish a standardized representation for each entity across its diverse applications. The centralized master repository architecture facilitates the creation of harmonized views of unique master objects, ensuring consistency and uniformity in data representation across the organization. Second, Management of Core Attributes: With a focus on managing data attributes consolidated from application data sets, the centralized master repository architecture aligns well with Van Gelder's objective of managing core master data objects within a single system. This approach allows for the establishment of a standardized set of core attributes associated with each master data model, providing a reliable source of "truth" for data management. Furthermore, Synchronization with Application Systems: While applications maintain their own local copies of master data, the centralized master repository serves as the centralized source of "truth." Synchronization mechanisms ensure that changes made to the master data set are propagated to all application systems, maintaining data consistency and integrity across the organization. Finally, Flexibility and Adaptability: The centralized approach offers flexibility in managing master data objects, allowing for different degrees of synchronization across application architectures. This adaptability enables Van Gelder to tailor the architecture to its specific business requirements, creating a good balance between harmonized views of master objects and the synchronization needs of individual applications.

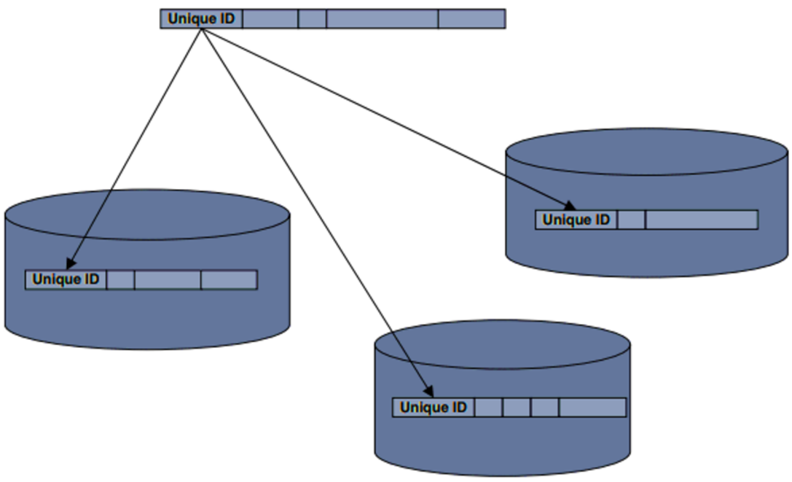


Figure 6.14: Centralized master repository architecture

6.3.2 Management guidance for MDM

Loshin has described a detailed guidance framework for the implementation of successful MDM [70]. As mentioned earlier, the implementation of MDM has been defined as a high-level action plan according to the framework of EA4DT, but Loshin said that the implementation of MDM has to be considered as a program rather than an individual project [68]. This is true and also the reason why the MDM project has been drawn as an overarching 'project' for the implementation of the PQi-cycle (fig. 4.6). For this reason, MDM can not be treated as an individual project, which can be completed after the completion of the real-time process control system project, but has to be included as a program and needs to be further developed over time and during the implementation of the remaining high-level action plan projects. This section of the report will briefly cover the steps that have to be taken to implement MDM in an organization, according to Loshin.

The first activity for implementation of MDM is Establishing a Business Justification. Loshin mentions that the implementation of MDM is coupled with other programs that benefit from it. For Van Gelder, these are mostly connected to the reporting standardization and the structured evaluating process, but also, the real-time process control system and predictive analysis model can benefit from the implementation of MDM. Besides the identification of the key, there are several more steps needed for this stage; these are listed below:

- Identification of key business drivers for master data integration
- Business impacts and risks of not employing an MDM environment
- Cost estimates for implementation
- Success criteria
- Metrics for measuring success
- Expected time to deploy
- Break-even point
- Education, training, and/or messaging material

The following step is to develop an MDM road map and roll-out plan. The MDM road map consists of the following steps: (1) Evaluate the business goals (2) Evaluate the business needs, (2) Assess the current state, (4) Assess the initial gap, (5) Envision the desired state, (6) Analyze the capability gap, (7) Perform capability mapping, and (8) Plan the project [70]. These steps are somewhat similar to the architecture context (sec. 4.1) and architecture vision (sec. 4.2). Some steps might have to be rewritten for the MDM purpose, but many steps can be completed based on these sections.
The roll-out plan has several stages, which are listed below. Figure 6.3 shows a road map with a roll-out plan.

1. Preparation and adoption
2. Proof of concept
3. Initial deployment
4. Release 1.0
5. Transition
6. Operations

Table 6.3: Road map and roll-out plan

	Data	**Services**	**Policy**	**Integration**
Preparation and Adoption	• Vendor models	• Vendor services	• Metadata management	• Data profiling • Simple identity search and match
Proof of Concept	• Customized models • Data lifecycle • Canonical model for exchange	• Core application services • Data quality, matching	• Business rules • Role-based access control	• Complex identity search and match—batch • Master registry
Initial Deployment	• Master relational models • Hierarchy management	• Functional application services	• Workflow management • Policy server	• Complex identity search and match—inline • Master registry
Release 1.0	• Browsing/editing	• Wrappers and facades	• Audit • Compliance • Service level agreements • Data governance	• Batch consolidation • Propagation
Transition	• Retire legacy data	• Value-added services	• Audit • Compliance • Service level agreements	• Synchronization • Coherence
Operational	• Enterprise information management • Data as a service	• Parameterized configuration • Software as a service	• Business policy management	• Interface with all enterprise applications

The next step is to define the roles and responsibilities of the stakeholders related to the data. This can be done with a so-called "RACI" model, which states who is Responsible, accountable, consulted, and informed of each activity [70]. All the stakeholders involved have to do this. From this, the business process models and the usage scenarios are drafted using the Identifying Initial Data Sets for Master Integration. Here, the business models are defined at a granular level, and the corresponding applications are defined [42]. This is also (partly) done in section 6.2.4 and shown in figure 6.12. For the MDM, the processes and applications can be worked out in more detail. The next step is to define the data governance. This includes the management processes for data governance and the tools to facilitate actionable governance [70]. The key deliverables of this stage are listed below:

- Management guidance for data governance
- Hierarchical management structures for overseeing standards and protocols
- Processes for defining data expectations
- Processes for defining, implementing, and collecting quantifiable measurements
- Definitions of acceptability thresholds for monitored metrics
- Processes for defining and conforming to directives in SLAs
- Processes for operational data governance and stewardship
- Identification and acquisition of tools for monitoring and reporting compliance with governance directives

The following step is metadata management, which acts as the backbone of MDM activities. The metadata repository plays a vital role in controlling the development and maintenance of the master data environment [70]. The milestones and deliverables for this step are listed below:

- Business requirements for metadata management tools
- Evaluation and selection of metadata tools
- Processes for gaining consensus on business data definitions
- Processes for documenting data element usage and service usage

The master object analysis and modeling are the following steps for the integration of MDM [70]. The master object analysis is responsible for determining which data is regarded as master data and requires the following results:

- Cross-system data analytic profiles
- Completed metadata profiles
- Business Rules
- Source-to-target mappings
- Extraction and transformation directives

Master object modeling emphasizes the development of models used for extraction of data and consolidation and sharing. For the process control system project, this is partly highlighted in chapter 6.
The next step is the data quality management. This is an essential step since the data quality determines the level of trustworthiness of the unified view [70]. The data governance guidance is used with metadata management and data quality techniques. The key deliverables for this part are listed below:

- The determination of business requirements for data quality tools
- The formalization and implementation of operational data quality assurance techniques
- The formalization and implementation of operational data governance procedures
- The acquisition and deployment of data quality tools

From the master object analysis and the data models, the data consolidation processes can be formulated. This phase involves the development of data extraction services. This phase is responsible for bringing data from different sources to one master data repository [70]. This is also explained for this project in section 6.2.4 and presented in figure 6.11. This phase has the following milestones and deliverables:

- Identification of business requirements for data integration tools
- Acquisition of data integration tools
- Acquisition of data consolidation tools
- Identification and business requirements for data consolidation technology, including identity resolution, entity search and match, and data survivorship
- Development of data extraction "connectors" for the candidate source data systems
- Development of rules for consolidation
- Implementation of consolidation services (batch, inline)
- Development of rules for survivorship
- Implementation of rules for survivorship
- Formalization of the components within the master data services layer

With all this, the MDM architecture style can be selected. The architecture style has already been chosen for this project in the section above (sec. 6.3.1). For this use case, the centralized master repository model is selected.
From this, the next step is to determine the services of the master data. These services can be put into three layers: (1) core service, (2) object services, and (3) business services [70]. The services of the process control system project have already been defined in chapter 5. The services for the master data can be implemented here as well or represented in a separate model. The key deliverables and milestones of this phase are listed below:

- Qualification of master service layers
- Design of master data services
- Implementation and deployment of master data services
- Application testing

The next step is to develop a transition plan. This plan is needed to transition to the newly created master data architecture and should help to mitigate the risk of transition failures [70]. This transition plan should be in line with the business constraint set in section 4.3.2.1. This phase consists of the following milestones and deliverables:

- Development of a migration/transition process
- Design and implementation of wrapper services
- Pilot transition for migration/transition process
- Institutionalization of migration/transition process

The final step is the ongoing maintenance. Maintenance is an integral part of the MDM architecture since the system should be reviewed to ensure the quality of the system [70]. It further helps mitigate existing or new applications in the architecture. The MDM manager(s) are responsible for continuous refinement of the architecture. The following expectations are set for this phase:

- Establishing minimum levels of service for data quality
- Establishing minimum levels of service for application performance
- Establishing minimum levels of service for synchronization
- Implementing service-level monitoring and reporting
- Enhancing governance programs to support MDM issue reporting, tracking, and resolution
- Reviewing ongoing vendor product performance

6.3.3 Decision-centric validation

As described in the research design (Chap. 3), the designed process control system will be validated by decision-centric validation, also known as Decision Centric Architecture Review (DCAR). This validation method requires less time and resources compared to other architecture validation methods like checklist-centric validation or scenario-based approach [71]. This makes it highly suitable for this research.
A vital aspect of the DCAR method is the analysis of the decisions made in the design of the architecture and the related decision forces. These decision forces can be described as "any non-trivial influence on an architect seeking a solution to an architectural problem." [72].
The DCAR methodology prescribes scheduling a half-day meeting and selecting around ten critical decision steps of the architecture to discuss within that session. For this research, it was not possible to schedule a half-day meeting with all the involved managers and content-related experts. Furthermore, the goal is to validate the entire architecture and not focus on just a selection of the decision choices. The reason for this is to give the opportunity to have discussions on topics related to architectural design that are not considered by the architect.
In total, eight invitations were sent for the decision-centric validation session, namely quality manager, operational manager, director, Health, Safety, and Environment (HSE) coordinator, process manager, data scientist, project manager IT, and an executor who has experience with the practical use of process control systems. Five of them were able to attend the meeting, but only the director, data scientist, and process manager were unable to join due to other obligations in their schedule. A week before the session, the report, including chapters 1-6, was sent to all attending members to give them the opportunity to read the report and ask for an explanation if something was unclear. The session started with an explanation of how the session was structured and what was expected of them during this session. The session was structured into four main parts. The first part served as background information, which is described in chapter 1 and 3. The second part covered chapter 5 including the architecture context, architecture vision, action plan, and outline. The third part covered chapter 5 including the conceptual architecture. Finally, the fourth part covered chapter 6 including the logical architecture and the data management. For each part, the report was used to go over each section and asked if there were any questions or comments on the decision making in the design of the system. Overall, the feedback from the participants was positive regarding the presented architecture. The architecture is regarded as sufficient to satisfy the predefined requirements, although several questions were raised to clarify some minor details, for example, how the new architecture changed compared to the current architecture. In the final report, these questions were addressed by providing more detailed explanations and elaborations. From this, it can be concluded that the system validation was successful due to a general consensus of the system's design.

7
Future Proofing

Before validating the current architecture, we will look at the future-proofing of the enterprise architecture. The first question that comes to mind is what is meant by future-proofing. The Cambridge Dictionary defines future-proofing as follows: *"to design software, a computer, etc. so that it can still be used in the future, even when technology changes"* [73] Rehman et al. defined future-proofing as: *"Future-proofing is the process of anticipating the future and developing methods of minimizing the effect of shocks and stresses of the future event"* [74]. This is the best fitting definition since it describes the word "*future*" as the process of anticipation and "*proof*" as developing mitigating strategies (to avoid obsolescence) [74]. Rehman et al. give three types of obsolete: functional, physical, and emotional. Here, functional obsolete means that the system functions less efficiently compared to other available products. Physical obsolete means that the system or product is deteriorated or worn out. Emotional obsolete means that the system or product is not desirable anymore [74]. The focus should be on functional and physical obsolesce for future-proofing since functional capabilities are the most essential part of the system. Furthermore, these obsolesce are quantifiable, whereas emotional obsolesce is subjective and much more complex to account for [74].

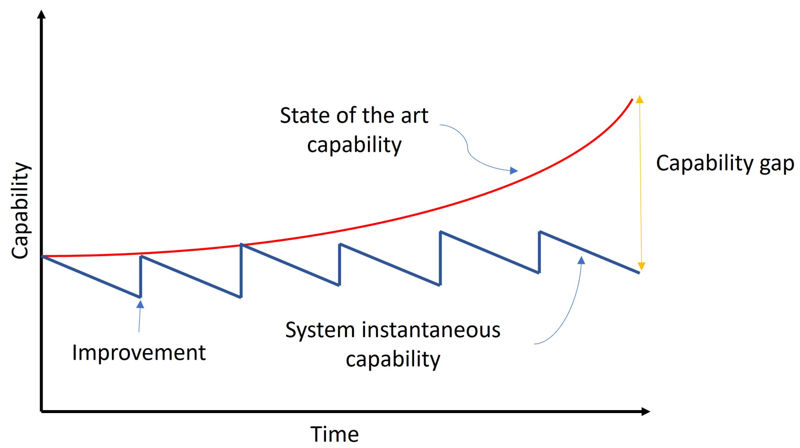


Figure 7.1: Capability gap of a system over time

One might say that system Life-Cycle Value (LCV) also could be a measure of the system's value in the future. LCV is, however, subjective since stakeholders are asked if the system meets the requirements. This makes the measure for the systems endurance less reliable [74]. The concept of future-proofing, on the other hand, is more related to the capability gap, where future state-of-the-art capabilities are compared to the current capabilities of the system, see figure 7.1.
Rehman et al. present a future-proofing process, which is shown in figure 7.2. Since the process control system is still in the conceptual phase, we are mostly interested in the left side of the figure. Here, it can be seen that the flow starts with the system requirements and continues to the predicted future requirements from which a predicted future solution is created. Due to time constraints, this research will not go through the entire process of the future-proofing process flow chart. The goal of this chapter is to define the predicted future requirements and propose predicted future solutions to meet these requirements.

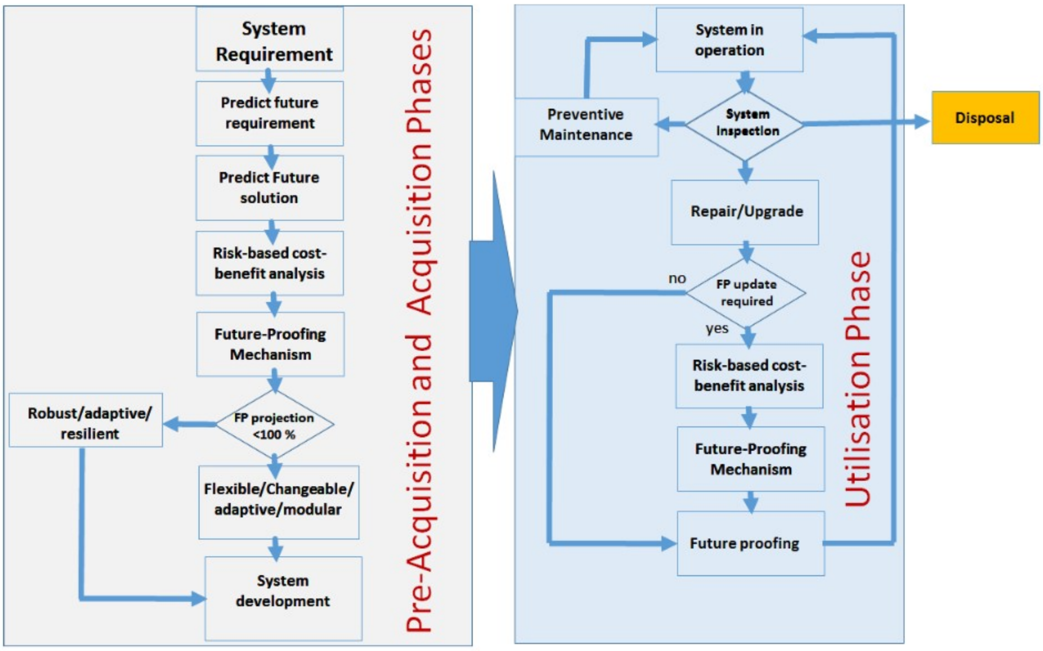


Figure 7.2: Future proofing process flow chart

7.1 Future requirements

Table 7.1 shows the current requirements and the expected requirement changes in the future. The changes are for the coming ten years and are based on genius forecasting. This method is based on basic logic and experience [74]. The expected future requirements also have not been quantified. The main goal is to find solutions that could accommodate the expected changes in the requirements.

Specific requirements are considered inherent characteristics of the process control system. These requirements are fundamental qualities of the system that do not vary in degree. Requirement numbers 1,2,5,6, and 8-10 are such inherent characteristics and are thus not expected to change in the coming ten years. The remainder of the requirements are more or less quantifiable and are expected to change in the coming years to increase the performance of the system. Requirements 3, 4, 7, 11, and 12 are all of a technical nature. The real-time data collection requirement could be increased to 1-5 seconds. The data storage and retrieval requirement will most likely be increased in the future due to expansions of the system or increased data quality, consequently expanding the data storage capacity. The alerts and notifications are now sketched as 2d visualizations on a screen, but the way this information is shown will most likely change in order to accommodate the additional information or make the interpretation easier. For example, 3D visualizations can be used, or better graphical design can be used to increase spatial awareness. The resolution requirement is probably also going to increase in the coming years. The expected cell size is thought to be around ten by 10 cm. Finally, the data quality is an important performance indicator of the system and will be increased to an error margin of around 5%.

Table 7.1: Current requirements and expected future requirements

nr.	Requirement	Future requirement (change)
1	Consistent data collection across diverse projects	/
2	flexibility for customization	/
3	Real-time Data Collection Requirement	Faster [1-5 seconds]
4	Data Storage and Retrieval Requirement	More storage capacity
5	Data Integration Requirement	/
6	User Access and Interface Requirement	/
7	Alerting and Notification Requirement	Better visualisations
8	Data Security and Compliance Requirement	/
9	Scalability and Performance Requirement	/
10	Reliability and Redundancy Requirement	/
11	Resolution requirement	Higher [10x10 cm]
12	Data quality requirement	Higher [5% error margin]

7.2 Solutions

This section will present several solutions to the expected future requirements set in the section above and represented in table 7.1. To define solutions for each requirement, this research will use a literature review to search for possible solutions for each requirement with an expected future change based on future trends. For each requirement 3,4,7,11, and 12 two future solutions we identified and listed below.

Real-time Data Collection Requirement - Requirement 3

- **Augmented Reality (AR) for Data Collection:** AR devices like smart glasses can provide real-time overlays of construction site data. Workers can use AR to view real-time data from sensors directly in their field of vision, improving situational awareness and data accuracy.
- **Advanced Wireless Sensor Networks (WSNs):** Implementing next-generation wireless sensor networks, such as those utilizing 5G technology, can facilitate faster and more reliable data transmission from numerous sensors deployed across the construction site.

Data Storage and Retrieval Requirement - Requirement 4

- **Distributed Ledger Technology (DLT):** Utilizing blockchain or other DLTs can provide secure, decentralized data storage solutions. This can ensure data integrity and quick retrieval while being scalable to accommodate growing data volumes.
- **Holographic Data Storage:** As an emerging technology, holographic data storage uses laser beams to store data in three dimensions within a medium. This can potentially offer high-density storage and quick data retrieval, suitable for the large amounts of data generated in construction projects.

Alerting and Notification Requirement - Requirement 7

- **IoT-enabled Wearables:** Wearable devices equipped with IoT capabilities can provide real-time alerts to workers and managers. These wearables can vibrate or emit sound alerts in response to data anomalies detected by the system.
- **Predictive Analytics Tools:** Advanced predictive analytics tools can analyze patterns in the collected data to forecast potential issues before they occur. By integrating machine learning, the system can send preemptive notifications to prevent problems.

Resolution Requirement - Requirement 11

- **High-Precision LIDAR:** Utilizing LIDAR technology can enhance spatial resolution and accuracy in data collection. LIDAR can create detailed 3D maps of construction sites, allowing for precise monitoring of construction activities.
- **Drones with High-Resolution Cameras:** Deploying drones equipped with high-resolution cameras can capture detailed images and videos of construction sites. This data can be processed to provide high-resolution spatial information.

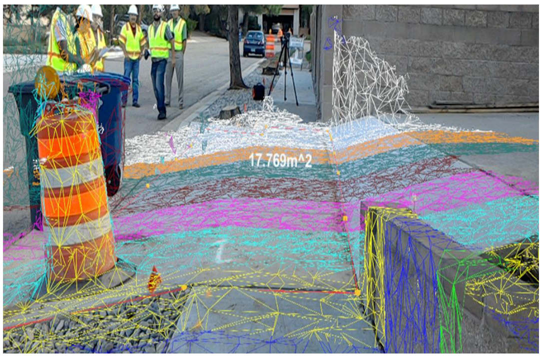
Data Quality Requirement - Requirement 12

- **Advanced Camera Systems:** Replacing GPS with advanced camera systems for tracking and data collection can improve accuracy. These cameras, equipped with AI, can analyze visual data to ensure quality and detect real-time errors.
- **Machine Learning for Data Cleaning:** Implementing machine learning algorithms specifically designed for data cleaning can automatically identify and correct errors in the collected data. These algorithms can continuously improve by learning from new data.

The first solutions for the real-time data collection requirement are Augmented Reality (AR) and Advanced Wireless Sensor Networks (WSNs). AR will most likely not increase the return rate of the data; on the contrary, it will most likely be slower since getting the visualizations as 2D on a tablet screen requires far less computational power. AR could serve as a solution for requirement 7, but this will be discussed later. WSNs is something already discussed in section 6.2.3. 5G is suggested for the future for faster and more reliable data transmission. This would indeed serve as a good alternative for 4G in the future if the requirements are set to higher standards and the coverage in the Netherlands is sufficient.

For the data storage and retrieval requirement, Distributed Ledger Technology (DLT) and holographic data storage are potential solutions for the expected future requirement to have more data storage and better data retrieval. DLT works based on the consensus protocol; here, changes made to the data are validated by the system without the need for a central governing body [75]. Holographic data storage, on the other hand, uses a technique to create more data storage and potentially better data retrieval. Both options are definitely overkill for the process control system for a single company. DLT is too complex for the current system and would only be suitable if many people are making changes to the system, which need to be validated first. The increase in data storage is something that is within the limit of current capabilities and could be easily solved by upgrading the cloud storage capacity and/or the capacity of the data hub.

For the alerting and notification requirement, IoT-enabled wearables and predictive analytic tools are potential future solutions. IoT-enabled wearable sensors could be a perfect solution for getting notifications or information to the receiving person. Instead of using a phone or tablet screen for notifications, a wearable item can give notifications or visualize the right data. This is where AR could play a vital role. AR glasses or a headset could be used to display the information for the operators to lower the cognitive load further since they do not have to focus on the work and a separate screen. Also, for site managers, this could be a solution to get a better understanding of the situation because spots are directly displayed on the road. The downside to AR is the safety concerns. The use of a headset might reduce situational awareness and thus pose a huge safety risk. This has to be thoroughly tested before implementing such a solution is possible. The other proposed solution is predictive analytic tools. This was also mentioned as a high-level action plan (sec. 4.3), where a model could be used to predict the outcome of future projects. Here, however, it could be used to produce better notifications promptly based on historical data. This could help prevent even more mistakes and would be very suitable for this project.



(a) Augmented Reality projection

(b) Augmented Reality headset

Figure 7.3: Augmented Reality

van gelder

The resolution requirement is now set at 25 by 25 centimeters; in the future, this is expected to go to around five by 5 centimeters. high-precision LIDAR and drones with high-resolution camera systems could realize this. The main issue with the lidar option is the fact that it is not capable of providing temperature data. Drones with high-resolution cameras could be a solution if the cameras are more advanced as they are now and able to shoot thermal footage. However, this would still be less favorable than mounting it on the back of the paver. This reduces the distance between the road and the camera, is more static, and is less susceptible to obstacles like trees blocking the view and being a danger for crashing the drone. In conclusion, neither option is suitable for getting a better resolution. A better option is to wait for higher-resolution cameras to replace the current ones.
For the final requirement, advanced camera systems and machine learning could be used for data cleaning. Here, the advanced cameras could replace the GPS sensors to improve accuracy and simultaneously analyze the visual data, which can serve as input for the predictive analytic tools, as mentioned earlier. Machine learning for data cleaning could be used to increase the quality of the recorded data by automatically identifying errors and correcting them. Another method is upgrading the calibration methods for the sensors, this can increase the data quality while the same equipment is used. Another possibility, like the thermal camera, is to upgrade existing sensors for more accurate ones if available. To conclude, a wide range of solutions were found; some are overkill or unfeasible, and others could be great solutions. It is essential to keep the design of the system as straightforward as possible while keeping the functional requirements in place. In this section, we did not look at additional functional requirements. It could, for example, be that further information is needed or an existing functional requirement is changed. It is essential to keep a close eye on both the functional -, and technical requirements in order to find potential (better) solutions.

8
Conclusion

The research objective of this research was to establish an enterprise architecture for a process control system that still allows Van Gelder to have the ability to make changes to the system, with the following research question: *How to establish an Enterprise Architecture framework within a road construction company to accommodate a real-time process control system that ensures consistent data collection across diverse projects (both inter and intra projects), facilitates flexibility for customization, and promotes future-proofing in the absence of ready-to-use solutions in the market?* In order to develop this enterprise architecture, the framework of Rozo was used. His paper presented a simplified framework for the successful implementation of an enterprise architecture tailored for digital transformation. The research question was split up into four separate research objectives. The following sections revisit and summarize the findings for each research objective in order to answer the main research question.

8.1 RO 1: Requirement definition

The first objective of this research was to define the requirements of the process control system. The research first established a basis by constructing the architectural context to determine the requirements. Here, the scope of the research was related to the two main objectives of Van Gelder: to generate trends, make predictions, and gather real-time process information. It was stated that the implementation of the PQi cycle was needed to realize these objectives. Empowering data-driven culture, Data management and governance, and Flexible and scalable infrastructure are the main strategies used to accomplish these objectives. Five new objectives were established more tailored to the needs of the PQi cycle implementation from the two objectives mentioned above. These are to improve the feedback, improve data collection, improve forecasting, Improve the effectiveness of suggested improvements, and have reasonable Quality assurance. For the architecture vision, the principles were defined, and these guided the decision-making for the remainder of the research. The capabilities were described in a value stream to give a clear overview of what was needed to master the PQi cycle. It was concluded that Van Gelder did not possess these capabilities as envisioned. The next step in the research defined the baseline and target architecture for the business level. Here, it was concluded that just three new functions were needed for a successful implementation of the PQi cycle: a data analyst to analyze the data, a process manager to manage the process control system, and the remainder of the PQi cycle, and a data engineer to maintain the data infrastructure. From this, the research defined five high-level action plans needed to implement the PQi cycle. It was concluded that each action plan could be executed individually, except for the master data management, since this was an overarching project. The architecture outline was selected as the high-level action plan project for the remainder of this research, for which all the technical and functional requirements were established in the architecture outline. Finally, it can be concluded that the research successfully established a strong foundation guided by the EA4DT framework and defined the requirements for the process control system.

8.2 RO 2: Develop a Data Infrastructure

The following research objective was to develop a data infrastructure using the framework for EA4DT. First, a conceptual architecture had to be developed. This research concluded that four services were needed to make the process control system functional within a company. A real-time information service is required to acquire the information and generate visualizations and notifications. Furthermore, an analysis and reporting access service is needed to give access to the data to the authorized personnel to analyze the data. Finally, data storage management is designed to account for the overarching data management mentioned earlier. Next, the data ontology was designed. The research concluded that the ontological model design used by Makarov could serve as the backbone for the new design, with additional attributes from literature that could make the model more complete. The research concluded that three types of classes needed to be defined, namely, static, quasi-static, and dynamic, and that the system's design had to take the characteristics of these classes into account during the design.

8.3 RO 3: Design a customizable process control system

For this objective, the research went into more detail about the logical architecture and analyzed two possibilities for processing the project data for real-time notifications and visualizations. The research concluded that cloud computing (central processing) was more desirable for the process control system. For data storage, it was concluded that the data needed to be stored temporarily on a data lake in the cloud and locally on the sensors for redundancy purposes. The data from the cloud data lake can be uploaded to Van Gelder's data hub. For the data transport, several options were explored and compared; the research concluded that 4G is currently the best option. To complete the logical architecture, the research presented the complete revised architecture, including the business -, application -, and technology levels. Finally, it was concluded that a centralized master repository architecture is needed to manage the data based on the number of attributes, the degree of consolidation, and the tightness of the coupling between the data sources. The research highlights the need for data management and found that it has to be included in the process control system project. A decision-centric validation method concluded that no comments and a general consensus validated the enterprise architecture design.

8.4 RO 4: Ensure Future-Proofing of the Process Control System

The final objective of the research was to future-proof the system in order to counter the absence of other ready-to-use solutions on the market and reduce the obsolescence of the system in the future. For this, the research found that first, the future requirement had to be estimated. A distinction was made between functional and technical requirements. For the technical requirements, changes were applied based on estimations to make them more suitable for future usage and system expectations. AI was used to answer the following question: *"Based on the provided architectural outline, the specific requirements listed (3, 4, 7, 11, and 12), and the technology architecture, what potential solutions can be proposed for these requirements, considering the anticipated technological advancements and trends over the next decade from which this system would benefit?"*. For each requirement, the AI bot generated two possible solutions. All the solutions were reviewed. Several proposed solutions were either unfeasible or would not solve the requirement change. For example, AI proposed AR to increase the real-time requirement and holographic data storage for higher-density data collection. However, several proposed solutions could work. For example, wearable IoT sensors for the notification requirement or cameras replacing GPS for more accurate location data. The research concluded that it is important to keep a close eye on both the functional -, and technical requirements in order to find potential (better) solutions.

Overall, it can be concluded that the research was successful in designing an enterprise architecture accounting for the complexities related to construction work by ensuring consistent data collection across diverse projects (both inter and intra-projects), facilitating flexibility for customization, see figure 6.12. Additionally, it provides several solutions to future-proof the system.

9 Discussion and future work

Although the enterprise architecture design has been validated, there are several discussion points. Initially, the architecture was written for Van Gelder, but the architecture design could also be used for other construction companies. The strategies and principles would stay the same as they form the foundation of the enterprise architecture (sec. 4.2.1). Still, there might be a need for slight modifications in the design. The business level, for example, could be entirely different within another organization. Despite this, the design can easily be projected at another road construction company with little work to tailor the enterprise architecture design. The business level for Van Gelder in the current design (fig. 5.4) might also need to be adjusted for the workload of the data engineer, data analyst, and process manager. The FullTime-Equivalent (FTE) for each business role needs to be assessed. In the beginning phase of the project, there will most likely be work for multiple data analysts. The structure of the business-level design would not change, but the functionalities of each data analyst might change due to this. In section 6.1, the data ontology was developed. The ontology was developed based on two ontologies from literature to have one complete ontology. Some attributes, however, are missing. For example, the underground temperature and the roller status are essential attributes that are currently not in the ontology. These, and maybe other attributes, might have to be included in the ontology based on future system testing. Furthermore, the asphalt core temperature station is included in the ontology, even though a model could be used to simplify the system. The reason for this is the quality of the (suggested) data. In the future, the station could be replaced with a model if enough data is available. This would be a separate project, simplifying the system, but only if the model is accurate enough to replace the station. The model also needs to be made for different asphalt types. In the technology level in figure 6.11, the technology level is highlighted. In this architecture, Metacom and PIM are connected via API to the data lake from which they are sent to the data hub of Van Gelder. Another way of gathering the data is using the current API, which directly connects PIM and Metacom to the data hub. For the continuation of this project, the best API route needs to be established. In the current architecture, the (predictive) analysis is done using a computer (local server). Depending on the amount of information and the analysis implementation, the computer could be upgraded to the cloud. This cloud increases the algorithm's performance if more capacity is needed. A computer would probably have enough capacity and performance for the implementation phase to run the algorithms. This brings us to another discussion point: the security risks of having corporate-sensitive information on the cloud. Even though data security is outside the scope of this project; it is essential to mention that this potential risk is clearly documented and mitigated before implementing the system. For future proofing, Artificial Intelligence (AI) was used to come up with possible solutions for future requirements. More time and careful analysis are needed to look at future trends and develop more solutions. Furthermore, only the technical requirements were changed for the future system, while functional requirements might also be changed based on changes in upcoming regulations. Finally, for future-proofing, the flow chart has to be completed; see figure 7.2. This would ensure that reasonable measures are in place to prove the system in the future. A final discussion point is the decision-centric validation. The architecture was validated since there were no comments on the design. Although this is a positive sign, it is remarkable that there were no comments or criticism points at all. This could be because of the attendees' lack of knowledge of such an architecture. As mentioned in section 6.3.3, five out of eight invited were present during the validation. Each was given plenty of time to read the report and ask questions beforehand if needed, which neither of them did. All attendees were true experts in their field but unfamiliar with such academic reports and architectures. This might have been a bit overwhelming. The three invited who could not attend the meeting (the director, process and information manager, and the data scientist) would most likely be more familiar with such reports and architectures. Another explanation could be the duration of the session and the amount of discussion points in the session. One and a half hours might have been too short to go over the entire report and prevented the discussion from going into detail. For future changes in architecture, a new decision-centric validation session could be planned for half a day, with all eight essential participants (stakeholders) involved.

Bibliography

[1] Stjepan Lakusic. Making the asphalt paving process explicit - A fundamental step for quality improvement. *Journal of the Croatian Association of Civil Engineers*, 72(11):1031–1040, 12 2020.

[2] Mohammad Sadeghian. *Develop an Ontology for Pavement Lifecycle Management.* Engineering doctorate thesis, University of Twente, Enschede, June 2023. EngD Thesis in the Department of Civil Engineering.

[3] Denis Makarov, Faridaddin Vahdatikhaki, Seirgei Miller, Afshin Jamshidi, and André Dorée. A framework for real-time compaction guidance system based on compaction priority mapping. *Automation in Construction*, 129:103818, 9 2021.

[4] TNO (Nederlandse Organisatie voor Toegepast Natuurwetenschappelijk Onderzoek). Tno - veilige en duurzame leefomgeving - infrastructuur - asfalt, Accessed 2024-02-08.

[5] Xingyi Zhu, Shunjie Bai, Genping Xue, Jian Yang, Yongsheng Cai, Wei Hu, Xiaoyang Jia, and Baoshan Huang. Assessment of compaction quality of multi-layer pavement structure based on intelligent compaction technology. *Construction and Building Materials*, 161:316–329, 2 2018.

[6] Qinwu Xu, George K. Chang, Victor L. Gallivan, and Robert D. Horan. Influences of intelligent compaction uniformity on pavement performances of hot mix asphalt. *Construction and Building Materials*, 30:746–752, 5 2012.

[7] Denis Makarov, Seirgei Miller, Faridaddin Vahdatikhaki, and André Dorée. Comprehensive real-time pavement operation support system using machine-to-machine communication. *International Journal of Pavement Research and Technology*, 13(1):93–107, 1 2020.

[8] Denis Makarov, Faridaddin Vahdatikhaki, Seirgei Miller, Sajjad Mowlaei, and André Dorée. Real-time trajectory planning for asphalt compaction operator support. *Automation in Construction*, 155:105034, 11 2023.

[9] Denis Makarov. Developing a real-time process control system for asphalt paving and compaction. Technical report, University of Twente, Enschede, 2018.

[10] Frank Bijleveld. *Asphalt road construction from an operational perspective.* PhD thesis, University of Twente, Enschede, The Netherlands, 1 2015.

[11] Arif Chowdhury and Joe W. Button. A Review of Warm Mix Asphalt . Technical report, Texas Transportation Institute, Texas A&M University System College Station, Texas 77843-3135, Texas, 12 2008.

[12] ASPARi - Asphalt Process Quality Improvement, 2023. Accessed: 8 January 2024.

[13] ASPARi. ASphalt PAving, Research & innovation (ASPARi): Towards professionalising the asphalt paving process, 2023.

[14] Yuan Ma, Feng Chen, Tao Ma, Xiaoming Huang, and Yang Zhang. Intelligent Compaction: An Improved Quality Monitoring and Control of Asphalt Pavement Construction Technology. *IEEE Transactions on Intelligent Transportation Systems*, 23(9):14875–14882, 9 2022.

[15] Wirtgen Group. Smart Compaction, 2023.

[16] Volvo. COMPACT ASSIST WITH DENSITY DIRECT™, 2023.

[17] Bomag. Compaction control: digital, transparent and in real time., 2023.

[18] Trimble. Compaction Control for Soil Compactors, 2023.

[19] Topcon. Thermal Mapper, 2023.

[20] Topcon. 3d-mc-asphalt, 2023.

[21] Moba. Pave-IR, 9 2023.

[22] Moba. Moba mobile compaction assistant, 9 2023.

[23] leica. Leica Icon Compaction, 2023.

[24] Leica. Leica iCON pave asfalt, 2023.

[25] Caterpillar. CAT® COMPACTION CONTROL TECHNOLOGIES R For Asphalt and Soil, 8 2016.

[26] Ammann. ACEpro – Ammann Intelligent Compaction Expert Reduces Compaction Costs, 2023.

[27] Dynapac. Digitized Road Contruction, 8 2023.

[28] moba. Moba-Matic, 9 2023.

[29] Q-Point - Q-Pave, Accessed 2024-02-08.

[30] Q-Point - Q-Compaction, Accessed 2024-02-08.

[31] Sakai. Intelligent Compaction (IC) & SAKAI CCV Sensor, 2023.

[32] Topcon. C-53 Intelligent Compaction, 2023.

[33] Voelkel - PAVE Navigator Brochure, Accessed 2024-02-08.

[34] Voelkel - FDVK Navigator Brochure, Accessed 2024-02-08.

[35] Vögele. VÖGELE RoadScan, 2023.

[36] Voelkel - Organic Juices and Drinks, Accessed 2024-02-08.

[37] Q-Point - Q-Machines 2, Accessed 2024-02-08.

[38] Sheunopa Chalmers Musukutwa. *SAP Enterprise Architecture: A Blueprint for Executing Digital Transformation*. Apress, New York, NY, 2022. This work is subject to copyright. All rights are reserved by the Publisher.

[39] Daniel F. Rozo. An enterprise architecture framework for digital transformation. Master's thesis, University of Twente, Enschede, The Netherlands, August 2020. Master Thesis.

[40] Tim Weilkiens. Validation techniques for software architectures. In Tim Weilkiens, editor, *Systems Engineering with SysML/UML: Modeling, Analysis, Design*, pages 71–86. Elsevier, 2016.

[41] Intel Corporation. Enterprise architecture: Enabling digital transformation. Technical report, Intel Corporation, 2018.

[42] Viktor Mayer-Schönberger and Kenneth Cukier. *Big Data: A Revolution That Will Transform How We Live, Work, and Think*. Eamon Dolan/Houghton

Mifflin Harcourt, 2013. https://link.springer.com/content/pdf/10.1007/978-1-4614-4214-1.pdf.

[43] Thomas C. Redman. *Data Quality: The Field Guide*. Digital Press, 2001.

[44] Kenneth Crow. *Process Optimization*. CRC Press, 2001.

[45] John Smith. *Continuous Improvement Planning*. Wiley, 2011.

[46] John Smith. Quality metrics implementation. *J. Qual. Manag.*, 25(3):201–215, 2020.

[47] John Smith. Change management: Principles and practices. *International Journal of Change Management*, 5(1):24–35, 2016.

[48] Emily Jones. Stakeholder engagement: A comprehensive guide. *Journal of Stakeholder Management*, 12(3):145–162, 2018.

[49] John Smith. Data source identification: Strategies and techniques. *Journal of Data Management*, 25(2):78–92, 2020.

[50] Sarah Jones. Data integration planning. *Int. J. Data Sci.*, 15(4):312–328, 2021.

[51] The Open Group. Understanding and guiding the architecture, 2024.

[52] The Open Group. Agile architecture framework: Constraints.

[53] Oxford english dictionary. https://www.oed.com/dictionary/consistent_adj?tab=factsheet#8368781. Accessed March 25, 2024.

[54] flexible, adj. https://www.oed.com/search/dictionary/?scope=Entries&q=flexible. Accessed on March 25, 2024.

[55] customizable, adj. https://www.oed.com/search/dictionary/?scope=Entries&q=customizable. Accessed on March 25, 2024.

[56] What is an ontology? https://www.oxfordsemantic.tech/faqs/what-is-an-ontology. Accessed on March 25, 2024.

[57] Parallels. Cloud vs server: Which is best for your business?, 2022.

[58] Aztech IT. Cloud computing vs traditional it: Pros and cons, 2024.

[59] Richard T. Watson. Data processing architectures, 2019. Accessed on April 15, 2024.

[60] Alexander S. Gillis. Cloud storage sla (service-level agreement), 2022. Accessed on April 15, 2024.

[61] Amazon Web Services. What is an API? https://aws.amazon.com/what-is/api/, 2024. [Online; accessed April 15, 2024].

[62] IBM. Internet of Things, 2024. Accessed: 2024-04-15.

[63] KPN Whitepaper. Iot connectiviteit: De 7 belangrijkste aandachtspunten. Online, 2023. Accessed: 2024-04-29.

[64] Consumentenbond. Welke internetsnelheid heb ik nodig? Online, 2023. Accessed: 2024-04-29.

[65] Vodafone. Netwerken | internet of things. Online, 2024. Accessed: 2024-04-29.

[66] Fysic. Uitfasering 2g en 3g, 2024.

[67] Beste Mobiele Providers. Dekkingskaarten, 2024.

[68] David Loshin. *Master Data and Master Data Management*, chapter 1. Elsevier/-Morgan Kaufmann, 2009.

[69] David Loshin. *Master Data and Master Data Management*, chapter 9. Elsevier/-Morgan Kaufmann, 2009.

[70] David Loshin. *Master Data and Master Data Management*, chapter 13. Elsevier/-Morgan Kaufmann, 2009.

[71] Marc H. Schmid, Cesare Pautasso, and Olaf Zimmermann. Decision-centric architecture reviews. In Len Bass, Paul Clements, and Rick Kazman, editors, *Software Architecture in Practice*, chapter 6, pages 157–185. Morgan Kaufmann, 2017.

[72] Uwe van Heesch, Visa-Pekka Eloranta, Paris Avgeriou, Ivano Malavolta, and Michel R.V. Chaudron. Decision-centric architecture reviews. *IEEE Software*, 31(1):69–76, January/February 2014.

[73] Future-proof.

[74] O. U. Rehman, M. J. Ryan, and M. Efatmaneshnik. Future proofing process. *INCOSE International Symposium*, 27(1):921–934, 2017.

[75] U.S. Government Accountability Office. Science tech spotlight: Quantum computing. Technical Report GAO-19-704SP, U.S. Government Accountability Office, 2019. Accessed: 2024-05-22.