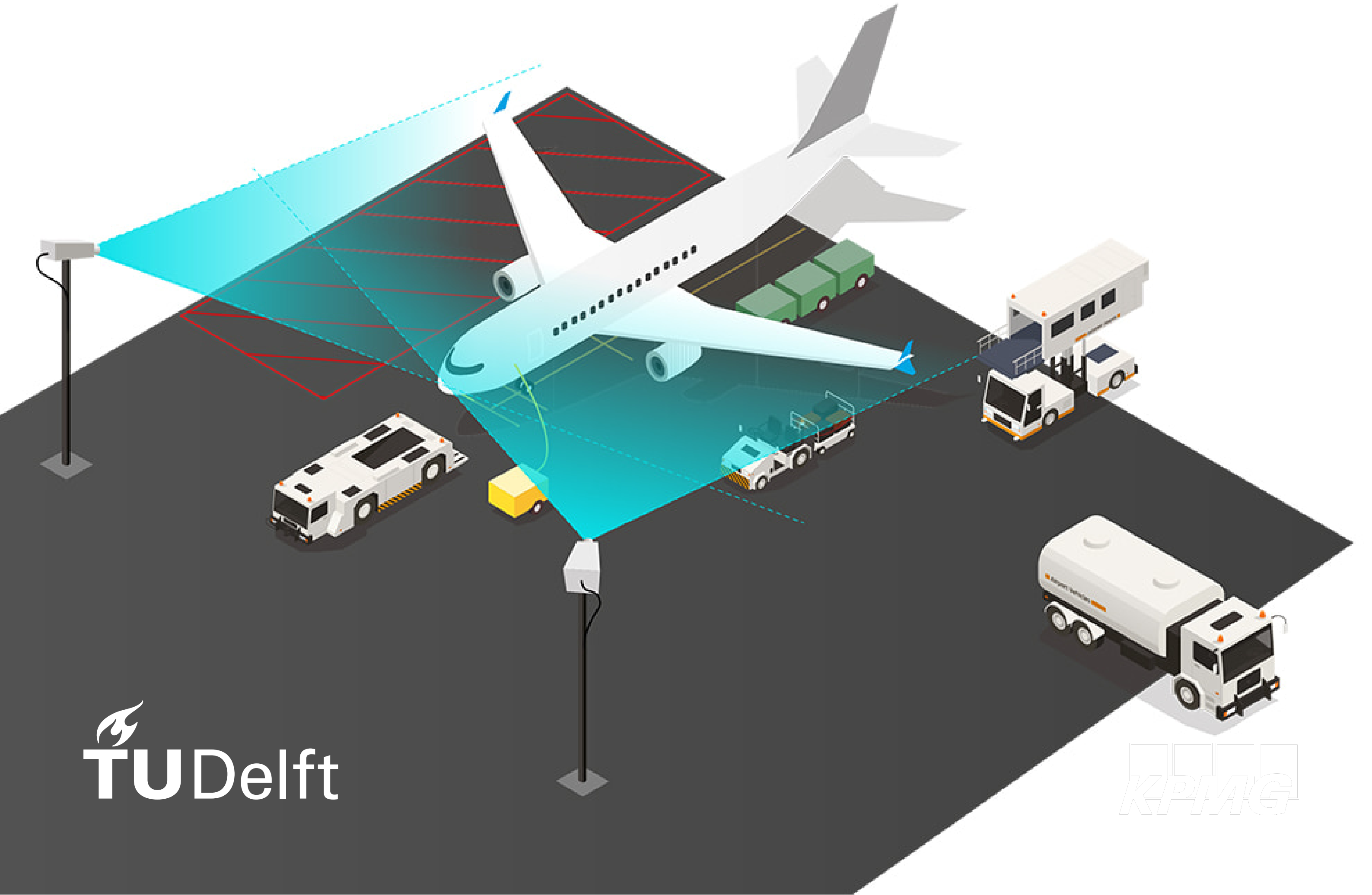
Flying Towards a Digital Twin

Designing a Digital Twin Maturity Assessment Framework for Airport Turnaround Operations

Master Thesis

Bodine Wilpshaar



Flying Towards a Digital Twin

Designing a Digital Twin Maturity Assessment Framework for Airport Turnaround Operations

by

Bodine Wilpshaar

to obtain the degree of Master of Science
at the Delft University of Technology,
to be defended publicly on Thursday July 2, 2026 at 2:15 PM.

Student number:	5302129	
Project duration:	March 19, 2026 – July 2, 2026	
Thesis committee:	Dr. J. A. Annema,	TU Delft, supervisor
	Dr. A. M. G. Zuiderwijk	TU Delft
	P. Groskamp	KPMG

Front cover image from Schiphol's blog post Turnaround Insights Schiphol, 2020.

An electronic version of this thesis is available at http://repository.tudelft.nl

Preface

Dear reader,

In front of you lies the report titled Flying Towards a Digital Twin: Designing a Digital Twin Maturity Assessment Framework for Airport Turnaround Operations. With the completion of this thesis, my academic career at TU Delft comes to an end after years of learning.

I would like to express my gratitude to the members of my graduation committee, who guided me academically throughout this final learning path. I am also grateful for the confidence they all had in taking on this topic with me, as we all learned more about it along the way.
Starting with Jan Anne: you always gave me the confidence I needed to trust myself, because you trusted me. Your academic eye was also extremely valuable. Anneke, your feedback, especially on the methodology, ultimately made me a better researcher and helped me take important steps forward in this process. It was also very special to have you both on my thesis committee, because Jan Anne was the first professor whose lecture I attended during my bachelor's at TU Delft, and Anneke was the last professor to hand me my bachelor's diploma. To end my master's programme in this way truly feels like coming full circle.
After a master's programme, the professional field comes into play. Here, I would like to thank KPMG for welcoming me into the team, as well as all my colleagues. I would especially like to thank Pieter, my thesis supervisor. During our meetings, he always had the ability to keep me on my toes and bring out the best in me, while doing so with a smile and in a supportive role. I also appreciated the conversations we had outside of the thesis.

Furthermore, I would like to thank all the people I interviewed. Especially my contact at Schiphol: without your help in setting up the meetings, this thesis would not be as valuable as it is. I would also like to thank the other three participants at Schiphol and the researcher who took the time to participate in multiple interviews. Of course, I am also grateful to the other participants I interviewed, who allowed me to ask every question that came to mind. Through the knowledge you shared, I experienced the best possible learning curve on a topic that was completely new to me.

And of course, my friends and family. My roommates made sure I could vent at the end of the day, brought me coffee and tea when I needed it most, and simply kept me company in my room, even on the floor. Thank you, Noa, for transcribing part of a topic you knew nothing about.
And thank you, Jasper, for always being there, for the flowers when I was at my most stressed, and for knowing what I needed when I sometimes had no idea myself.
And of course, my parents, brother and Donnie. As my parents said: "I am always team Bo" and I have never doubted that. Without your support, and without you making sure I could attend university in the first place, this would not have been possible.

Lastly, this thesis marks the end of my master's programme. But in the end, I could not have done this without my support team. So: GO TEAM!

B.A. Wilpshaar
Utrecht, June 24, 2026

Executive summary

Airports are under increasing pressure to use existing capacity more efficiently while maintaining operational reliability, safety, and service quality. A critical process in this context is the aircraft turnaround: the set of time-sensitive activities required to prepare an arriving aircraft for its next departure. Turnaround operations involve many interdependent activities and multiple stakeholders, including the airport operator, airlines, ground handlers, air traffic control, and network management. As a result, delays or information gaps in one part of the process can quickly affect departure punctuality, gate availability, resource planning, passenger connections, and wider airport performance.

Airport Collaborative Decision Making (A-CDM) is a way for airport stakeholders to coordinate operations by using shared procedures, milestones, and target times. However, A-CDM does not by itself create a fully integrated, continuously updated, and decision-supporting representation of the turnaround process. Digital twins offer a promising next step: in this research, a digital twin is understood as a digital representation of a physical process or system that is updated with operational data and can support monitoring, analysis, prediction, and decision-making. In the turnaround context, these capabilities depend on the maturity of the digital twin. A low-maturity digital twin may mainly show static or isolated information, while a high-maturity digital twin can support shared monitoring, prediction, and operational decision support. Existing maturity models provide useful foundations, but they are mainly developed for single-asset, infrastructure, manufacturing, building, or other technically bounded contexts. Airport turnaround operations require a more specific framework because maturity depends not only on the technical capabilities of the system, but also on coordination between independent actors with different systems, responsibilities, data interests, decision rights, and safety constraints.

This thesis addresses the following research question: Which maturity assessment framework can guide the development of digital twins in airport turnaround operations? The research follows a Design Science Research approach. It first explicates the problem and reviews literature on digital twins, airport turnaround operations, A-CDM, and maturity models. It then defines requirements, develops and refines the framework through expert feedback, demonstrates the framework in the Schiphol context, and evaluates its clarity, relevance, and potential usefulness through feedback from Schiphol and Vancouver International Airport.

The main outcome of this research is a digital twin maturity assessment framework for airport turnaround operations. The framework consists of a baseline category and five maturity levels: **Digital Model, Digital Shadow, Monitoring Twin, Predictive Twin, and Optimizing Twin**. These levels describe a development path from a static or non-real-time digital representation towards a digital twin that supports shared monitoring, prediction, and optimisation-oriented decision support. The highest level is deliberately not defined as full autonomy, because airport turnaround operations remain dependent on human judgement, organisational decision rights, workflow boundaries, accountability structures, and safety requirements.

The framework assesses maturity through six **internal maturity dimensions: Data Infrastructure, System Integration, Predictive and Analytical Capability, Operational Decision Optimisation, Operational Situation Visibility, and Human Interaction and Trust**. These dimensions make maturity measurable as a multidimensional construct rather than as one single overall score. Internal maturity concerns what the digital twin is technically and operationally capable of doing, such as capturing data, integrating systems, monitoring the current situation, predicting future states, supporting operational decisions, and presenting information in a usable and trustworthy way. The framework also distinguishes these internal capabilities from **external implementation constraints**. These constraints determine whether digital twin capabilities can actually be used in practice and include data access, ownership, privacy, commercial sensitivity, decision rights, accountability, workflow integration, and safety requirements.

The framework was first demonstrated at Amsterdam Airport Schiphol, where initiatives such as Wilbur and

Deep Turnaround provide a relevant case for assessing turnaround-related digital twin maturity. The purpose of the demonstration was not only to assign a maturity score, but also to examine whether the framework could be applied in a real airport turnaround context and whether it supported a useful discussion about current capabilities, limitations, and development priorities.

The Schiphol demonstration showed that Schiphol has a strong foundation in operational visibility, timely data availability, system integration, and selected predictive capabilities. At the same time, the assessment showed that maturity is uneven across dimensions. Important development areas include improving data completeness and data quality management, representing task dependencies and knock-on effects more explicitly, developing scenario analysis, and translating predictive insights into workflow-ready operational recommendations such as resource reassignment, task resequencing, and structured decision support. The Schiphol feedback indicated that the framework was understandable and useful as a structured discussion tool, because it helped participants discuss digital twin maturity across data, integration, prediction, decision support, operational visibility, and human interaction.

The framework was then evaluated with Vancouver International Airport as an exploratory transferability check. This evaluation did not aim to score YVR. Instead, it examined whether the framework was understandable, relevant, and meaningful in another airport context with a different digital twin development pathway. The YVR feedback indicated that the main structure of the framework was transferable beyond Schiphol. In particular, the maturity levels, internal dimensions, and distinction between internal maturity and external constraints were considered relevant. At the same time, the evaluation showed that the interpretation of specific scores may depend on local governance structures, data-sharing arrangements, operational responsibilities, and regulatory conditions. This means that the framework can be transferred as an assessment structure, but should not be used as a fixed scoring template for direct comparison between airports without contextual interpretation.

This research contributes both practically and academically. In practice, it provides airport operators with a structured instrument to assess current digital twin capabilities and identify targeted development steps. Rather than asking whether an airport "has" a digital twin, the framework helps determine what type of digital twin capability exists, how mature it is across different dimensions, and what is needed to move towards more integrated, predictive, and decision-supporting use. The Schiphol demonstration showed how the framework can be used to identify current strengths and improvement areas in a specific airport context, while the YVR evaluation suggested that the framework can also support structured reflection in other airport settings.

Academically, the thesis adapts general digital twin maturity thinking to the specific context of airport turnaround operations. It shows that maturity in this context depends not only on technical advancement, but also on shared situational awareness, multi-stakeholder coordination, trusted human interaction, and operational use within governance and safety constraints. The distinction between internal maturity and external implementation constraints contributes to a more socio-technical understanding of digital twin maturity, because it explains why technically advanced capabilities may still have limited operational use when data access, commercial sensitivity, decision rights, workflows, accountability, or safety boundaries restrict implementation.

The research should be interpreted as an initial framework development and evaluation. The framework is developed and demonstrated mainly in the Schiphol context, and the scoring logic is not yet a fully validated quantitative measurement instrument. The YVR evaluation provides an exploratory indication of transferability, but broader validation is still needed. Future research should apply the framework across multiple airports, repeat the assessment with different stakeholder groups, and further test the scoring logic quantitatively. Overall, the thesis concludes that digital twin maturity in airport turnaround operations should be assessed through maturity levels, internal dimensions, indicators, external implementation constraints, and structured scoring logic. In this way, the framework supports the gradual development of airport turnaround digital twins from digital representation towards shared monitoring, prediction, and human-governed operational optimisation.

List of abbreviations and terms

A-CDM	Airport Collaborative Decision Making
DSR	Design Science Research
Schiphol	Amsterdam Airport Schiphol
TOBT	Target Off-Block Time
YVR	Vancouver International Airport

Contents

Part A

Explicate Problem

1
Introduction

The aviation sector continues to grow, placing increasing pressure on airport capacity, operational coordination, and service quality. ACI World forecasts that global passenger traffic will reach 18.8 billion passengers by 2045 (ACI World, 2026). For large hub airports, this growth creates a need not only for additional infrastructure, but also for more efficient use of existing airport processes. As airport systems become busier, delays and disruptions can spread more easily across operations, making coordination increasingly important (Evler et al., 2022; Schmidt, 2017).

One of the most critical airport processes in this context is the aircraft turnaround. Turnaround refers to the activities required to prepare an arriving aircraft for its next departure, such as deboarding, baggage handling, cleaning, catering, fuelling, boarding, and operational checks (Schmidt, 2017; Wu & Caves, 2000). These activities are time-critical and strongly connected. A delay in one activity can influence other activities and may eventually delay the aircraft departure. Because airport operations are closely connected, turnaround delays can also affect gate availability, resource planning, passenger connections, and later flights in the wider airport network (Evler et al., 2022; Schmidt, 2017; Wu & Caves, 2000).

Traditionally, airport turnaround processes are often managed through separate systems and by different stakeholders, each focusing on its own part of the operation. This can make coordination difficult, because information about the same turnaround process may be spread across airport operators, airlines, ground handlers, air traffic control, and the network manager. To bring these separate processes and information flows closer together, some airports use Airport Collaborative Decision Making (A-CDM). A-CDM supports the timely sharing of operational information between these actors. Its purpose is to improve common situational awareness, predictability, and decision-making during airport operations, especially around turnaround and departure processes (Cappellazzo, 2025; EUROCONTROL, n.d.; Vail et al., 2015). A-CDM therefore provides an important basis for shared operational coordination.

However, A-CDM mainly provides a coordination framework for sharing operational information. It does not by itself guarantee that all turnaround data are integrated, continuously updated, or translated into decision support. In practice, stakeholders may still work with different systems, data formats, responsibilities, and access rights. As a result, creating one shared and reliable operational picture remains difficult. This means that the next challenge is not only to share information, but also to connect, interpret, and use that information across organisational and technical boundaries (EUROCONTROL, n.d.; Vail et al., 2015).

Digital twins can be seen as a promising next step to strengthen this coordination. A digital twin can be understood as a dynamically updated digital representation of a physical system that uses data to support monitoring, analysis, prediction, and decision-making (Singh et al., 2021; Van Der Horn & Mahadevan, 2021). In the airport turnaround context, a digital twin could connect data from multiple stakeholders, create a shared operational picture, identify emerging disruptions, predict delay effects, and support more proactive decision-making.

This support is only possible when the digital twin has reached a sufficient level of maturity. Digital twin maturity refers to the extent to which a digital twin has developed from a basic digital representation towards more advanced capabilities such as real-time monitoring, prediction, decision support, and optimisation (L. Chen et al., 2021; Li et al., 2024; Uhlenkamp et al., 2022). A low-maturity digital twin may only show static or isolated information and therefore cannot fully support collaborative decision-making. This is problematic because airports are encouraged to use real-time operational data, IT, and AI tools to monitor aircraft turnaround, queues, and delays (ACI EUROPE, 2025). Since maturity models are used to assess current capabilities and guide further development, a maturity assessment is needed to determine whether these digital tools can actually support operational use. Without such an assessment, airports may invest in digital tools without knowing whether the required capabilities are actually present.

Existing digital twin maturity models provide useful concepts for assessing these capabilities. However, many maturity models are developed for more clearly bounded systems, such as machines, assets, buildings, infrastructure components, or manufacturing systems (L. Chen et al., 2021; Li et al., 2024; Uhlenkamp et al., 2022) (chapter 3). Airport turnaround operations are different. They are short-cycle, safety-critical, highly time-dependent, and depend on coordination between multiple independent actors. Therefore, the maturity of a turnaround digital twin cannot be assessed only as a technical development path towards prediction and optimisation. It must also consider whether the twin can support multi-stakeholder data integration and shared situational awareness.

This creates both a practical and an academic gap. The practical gap is that airports need a structured way to assess whether digital twin technologies are mature enough to support collaborative turnaround decision-making. The academic gap is that existing digital twin maturity models do not fully address the multi-actor and coordination-intensive nature of airport turnaround operations. As a result, these models cannot be directly applied to the turnaround context without adaptation.

This study addresses this gap by developing a digital twin maturity assessment framework for airport turnaround operations. The aim is to evaluate how mature current digital twin capabilities are and to identify how further development can be guided. The detailed problem definition and main research question are presented in section 1.1.

1.1. Research gap and main research question

Digital twins can support airport turnaround operations by improving monitoring, prediction, decision-making, and shared situational awareness. However, these benefits depend on the maturity of the digital twin. A low-maturity digital twin may only show static or isolated information and may therefore not be able to support collaborative decision-making in a complex turnaround environment.

Based on this practical and academic gap, this research focuses on the lack of a maturity assessment framework that is specifically designed for digital twins in airport turnaround operations. Such a framework is needed to assess current capabilities and guide further development towards better coordination, shared situational awareness, prediction, and decision support.

Therefore, airports lack a maturity assessment framework that is specifically designed for digital twins in airport turnaround operations. Such a framework is needed to assess current capabilities and guide further development towards better coordination, shared situational awareness, and decision support.

This leads to the following main research question:

Which maturity assessment framework can guide the development of digital twins in airport turnaround operations?

The research methodology and the sub-questions that structure the study are presented in chapter 2.

1.2. Research scope and case selection

This study uses a case-based research design to refine, demonstrate, and evaluate the maturity assessment framework. This approach is suitable because digital twin maturity in airport turnaround operations is context-dependent and shaped by technology, operational processes, data availability, stakeholder coordination, and institutional constraints. The cases were selected purposively rather than randomly. They were chosen because they provide analytically relevant insight into the research problem, rather than being statistically representative of all airports (Flyvbjerg, 2006; Seawright & Gerring, 2008).

The research focuses on the airport level rather than on individual ground handlers. Although ground handlers execute many turnaround activities, they form only one part of the wider turnaround system. Turnaround operations depend on coordination between the airport operator, airlines, ground handlers, air traffic control, and other stakeholders. Since this research focuses on shared situational awareness and collaborative decision-making, the airport level is the most suitable level of analysis.

Amsterdam Airport Schiphol (Schiphol) is selected as the main case. Schiphol is an information-rich case because it combines a large and complex hub operation with A-CDM, Wilbur, and Deep Turnaround. These elements make Schiphol suitable for developing and demonstrating the framework in a turnaround context where digital-twin-related capabilities are already present and expected to be relatively advanced (Flyvbjerg, 2006; Seawright & Gerring, 2008).

Schiphol is not starting from zero. Since 2017, Schiphol has been developing Wilbur, a real-time operational platform that brings airport process data together, creates a shared "data truth", looks four to eight hours ahead, and gives machine-learning-based suggestions for expected constraints (Royal Schiphol Group, 2020). Schiphol also developed Deep Turnaround, a turnaround-specific solution that uses AI image processing to detect and report more than 70 turnaround events across 30 turnaround processes. It provides historic, real-time, and predictive insights (Royal Schiphol Group, 2024). In this thesis, Wilbur is understood as Schiphol's broader operational digital twin environment, while Deep Turnaround is understood as a specialised turnaround intelligence layer that feeds data to Wilbur.

Vancouver International Airport (YVR) is selected as a contrasting external evaluation case. This follows the logic of purposeful and contrasting case selection, in which cases are chosen because they provide relevant contextual variation rather than statistical representativeness (Seawright & Gerring, 2008). YVR differs from Schiphol because it is not used here as an A-CDM-based turnaround case. Instead, it represents an airport with an advanced airport-wide digital twin environment, while the turnaround-specific part of that digital twin is still developing.

This contrast is useful for evaluating the transferability of the framework. YVR has developed an airport digital twin that uses historical and real-time data to support data-driven decision-making, collaboration, and situational awareness (Unity, n.d.). At the same time, YVR has recently started implementing Deep Turnaround, the turnaround intelligence solution originally developed by Schiphol (Royal Schiphol Group, 2024, 2025). Therefore, Schiphol provides depth through a relatively advanced turnaround-focused case, while YVR provides variation by showing whether the maturity assessment framework is also understandable and relevant in an airport where the broader digital twin foundation is advanced, but the turnaround-specific maturity is still emerging.

1.3. Connection to CoSEM study

This research is carried out within the Master of Science programme in Complex Systems Engineering and Management (CoSEM) at Delft University of Technology. It fits within the Transport and Logistics domain because it focuses on airport operations and, more specifically, aircraft turnaround processes.

The topic fits CoSEM because airport turnaround operations are complex socio-technical systems. They depend on technical systems, physical infrastructure, operational processes, and collaboration between stakeholders such as airport operators, airlines, ground handlers, air traffic control, technology providers, and regulators. Digital twin development in this context is therefore not only a technical challenge, but also an organisational and institutional one.

This thesis combines these perspectives by developing a maturity assessment framework for digital twins in airport turnaround operations. The technical perspective focuses on digital twin capabilities and system development. The process perspective focuses on how these capabilities can support turnaround operations. The institutional perspective focuses on stakeholder collaboration, data sharing, governance, and decision-making.

The research has a clear design component. The framework helps airports assess their current digital twin maturity and identify possible next development steps. In this way, the study contributes to a complex socio-technical problem while considering both public and business values, such as operational reliability, passenger experience, data sharing, governance, and trust. This fits the CoSEM approach because the thesis does not only analyse a complex socio-technical problem, but also designs and evaluates a practical artefact to support decision-making in that system.

1.4. Thesis structure

This thesis is structured according to the main phases of Design Science Research. Chapter 1 introduces the research context, scope, case selection, research gap, and main research question. Chapter 2 explains the methodology and sub-questions. Chapter 3 presents the theoretical background on digital twins, airport turnaround operations, and maturity models. Chapter 4 defines the requirements for the framework. Chapters 5, 6, and 7 describe the development and refinement of the maturity assessment framework. Chapter 8 demonstrates the framework in the Schiphol case, and Chapter 9 evaluates its usefulness with feedback from Schiphol and YVR. Finally, Chapter 10 answers the main research question, interprets the findings, discusses the academic and practical contributions, and reflects on the limitations and directions for future research.

2 Method

Design Science Research (DSR) is used as the umbrella methodology for this study. The main research question asks which maturity assessment framework can guide the development of digital twins in airport turnaround operations. This makes the research design-oriented, because the aim is not only to analyse a problem, but also to develop and evaluate a practical artefact. This aligns with the design science paradigm, in which knowledge is generated through the creation and evaluation of artefacts that address relevant practical problems (Hevner et al., 2004). In this thesis, the artefact is the digital twin maturity assessment framework.

DSR was also chosen because it helps to make the development process of the framework transparent. This was important because maturity model development has been criticised for often providing limited documentation of the procedures and methods through which models are developed (Becker et al., 2009). In this thesis, the maturity assessment framework is not presented as a final outcome without explanation, but developed through a step-by-step design process. The DSR structure helps to show how the problem was first explicated, how requirements were defined, how these requirements were translated into maturity levels, dimensions, indicators, and scoring logic, and how expert feedback was used to refine the artefact. In this way, DSR is suitable not only because this thesis designs an artefact, but also because it makes the reasoning behind the framework more traceable and transparent (Hevner et al., 2004; Johannesson & Perjons, 2014).

The methodological foundation of the study is informed by the established Design Science Research literature. Hevner et al. (2004) provide the general logic and quality criteria of design science research, emphasising problem relevance, artefact creation, rigorous evaluation, and research contribution. Peffers et al. (2007) further formalise Design Science Research into a methodology consisting of problem identification and motivation, definition of solution objectives, design and development, demonstration, evaluation, and communication. These works are used to justify the suitability and scientific rigour of the research approach.

To structure the actual research process, this thesis follows the five-activity DSR framework of Johannesson and Perjons (2014). This framework is selected because it translates the broader DSR logic into a clear and traceable sequence of research activities. The five activities are:

A. Explicate Problem
B. Define Requirements
C. Design and Develop Artefact
D. Demonstrate Artefact
E. Evaluate Artefact

These five activities structure the thesis. The problem is explicated through the introduction and theoretical background. Requirements are then defined to determine what the framework should assess and under which airport-specific conditions it should be useful. The framework is subsequently designed and developed, demonstrated in the Schiphol case, and evaluated through expert feedback and questionnaire responses. In this way, the research approach combines the general DSR principles of Hevner et al. (2004) and Peffers et al. (2007) with the practical activity structure of Johannesson and Perjons (2014).

The choice for Design Science Research also has limitations. First, Hevner et al. (2004) emphasise that a design artefact should be rigorously evaluated. In this study, the framework is evaluated through a demonstration at Schiphol, a questionnaire, and expert feedback. However, because the framework is mainly demonstrated and evaluated in one airport context, the results should be interpreted as an initial validation rather than as proof of general applicability to all airports. This limitation is addressed by including expert input from different perspectives and by reflecting on the applicability of the framework beyond the Schiphol case.

Second, the DSR process requires several translation steps: from literature to requirements, from requirements to framework dimensions and indicators, and from expert feedback to design improvements. These steps involve researcher interpretation. To reduce this limitation, the design process is made traceable by documenting how literature, requirements, framework elements, and design decisions are connected.

Third, Johannesson and Perjons (2014) describe DSR as an iterative process in which design choices can be refined throughout the study. This fits the development of the framework, but it also means that the final artefact depends on the available literature, selected experts, and practical case context. Therefore, the framework is positioned as a structured assessment and reflection instrument for airport turnaround operations, rather than as a universal prescription for digital twin maturity in all airport settings.

2.1. Operationalisation of the DSR Activities

This section explains how the five DSR activities were operationalised in this study. For each activity, it describes the methodological purpose, the main research actions, and how the output informed the next step of the design process.

2.1.1. A. Explicate Problem

The first activity in the Design Science Research (DSR) process is *Explicate Problem*. This activity clarifies the practical and academic problem that motivates the design study. Johannesson and Perjons (2014) describe a practical problem as a gap between a current state and a more desirable state within a practice. In this research, the current state is that airports are developing digital-twin-related capabilities for turnaround operations, while a structured way to assess the maturity of these capabilities is still missing. The desired state is a maturity assessment framework that helps airports evaluate current capabilities and identify suitable next development steps.

The problem explication was conducted through a combination of practical problem framing and literature review. The practical problem framing was used to define why airport turnaround operations form a relevant design context. Turnaround operations are time-critical, multi-actor processes in which coordination, data exchange, and shared situational awareness are essential. This makes the topic suitable for an artefact-based approach, because the research aim is not only to describe the problem, but to design a framework that can support assessment and improvement in practice.

The academic problem was explicated through a structured review of literature on digital twins, airport turnaround operations, airport digital twin applications, and digital twin maturity models. The literature review followed a PRISMA 2020-inspired structure to make the search and selection process transparent. PRISMA 2020 supports the reporting of how studies are identified, screened, selected, and included (Page et al., 2021). In this study, it was used as a supporting review structure to document the search strings, database selection, inclusion and exclusion criteria, screening steps, and final set of selected papers.

This structured review approach also has limitations. A PRISMA-inspired review increases transparency, but it can be less suitable for emerging and practice-oriented fields where relevant insights may not yet be fully available in academic literature. A broader narrative review could therefore also have been useful. However, the structured approach was still considered appropriate because the review was used to create a clear and traceable academic basis for the framework. Industry and practitioner sources were only included in specific parts where they provided relevant contextual insight into airport digital twin applications or turnaround-related technologies.

PRISMA 2020 was therefore not used as a complete design method. It supported the transparency of the literature selection, but it did not determine how the selected literature was interpreted or translated into framework requirements and design elements. This interpretation remained part of the DSR process, in which the knowledge base is used to understand the problem, derive requirements, and construct the artefact.

The *Explicate Problem* activity was not translated into a separate sub-question. Instead, it provided the foundation for the subsequent design activities. The output of this activity was a clear problem definition, a justified research gap, and a literature base that could be used in the next steps. Chapter 3 therefore forms an important foundation for the later chapters, because the literature reviewed there informs the derivation of requirements in SQ1 and supports the development of maturity levels, dimensions, indicators, and assessment logic in SQ2.

2.1.2. B. Define Requirements

This DSR activity answers the following sub-question:

SQ1 - What requirements should a digital twin maturity assessment framework fullfil to support the development of a digital twin for airport turnaround operations?

The second activity in this research was the definition of requirements. This activity corresponds to the *Define Requirements* phase in the design science research method of Johannesson and Perjons (2014). In this phase, the conditions that the artefact should fulfil are made explicit before the artefact is designed. This was important in this research because the artefact is not a digital twin itself, but a maturity assessment framework for airport turnaround digital twins. Therefore, the requirements describe what the framework should assess, how it should be organised, and under which contextual conditions it should be applicable.

The need for a separate requirement-definition phase became clear during the literature review. Several existing digital twin maturity models provide useful levels, dimensions, indicators, rubrics, and scoring approaches, but they often do not make explicit how the requirements for the maturity model were first derived. For example, L. Chen et al. (2021) present a Gemini Principles-based model with three dimensions, nine sub-dimensions, and 27 rubrics. Uhlenkamp et al. (2022) develop a maturity model with seven categories and 31 ranked characteristics, based on a systematic literature review and use-case validation. Li et al. (2024) propose a five-level maturity hierarchy with five dimensions and 13 rubrics, supported by a D-ANP scoring method. These studies are valuable for maturity assessment, but they mainly present the final model structure and evaluation method. Therefore, this research adds a separate requirement-definition phase to make the link between the literature and the final framework more transparent.

Because of this, the present study added an explicit requirement phase before designing the maturity assessment framework. This makes the transition from literature to framework more transparent. Instead of directly selecting dimensions and indicators, the literature was first analysed to identify conditions that a useful maturity assessment framework should fulfil. These conditions were then grouped into requirement themes. This helped to make clear why each part of the framework was included and how it was supported by existing research.

The requirement definition followed three steps. First, the selected literature from the theoretical background (chapter 3) was revisited with a specific focus on statements, concepts, findings, and design choices that were relevant for maturity assessment. Because the literature did not always use the same wording, the selection was not based only on exact keywords. Instead, conceptually similar statements were grouped together. For example, statements about real-time monitoring, current operational state, and live data exchange were grouped under real-time operational visibility.

Second, the selected evidence was grouped into functional, structural, and environmental requirements. This categorisation follows Johannesson and Perjons (2014), who distinguish between requirements related to the functions of an artefact, its structural qualities, and its environmental qualities. In this research, functional requirements describe what the framework should assess in a turnaround-focused digital twin. Structural requirements describe how the maturity assessment framework itself should be organised. Environmental requirements describe the broader conditions under which the framework should be applicable.

Third, a traceability approach was added. For each requirement, the supporting papers were recorded in a traceability matrix. In addition, detailed evidence tables were created to show which statement from which source supported each requirement. This was done to address the lack of explicit traceability observed in several maturity model studies. The traceability tables make the derivation of the requirements more transparent and reduce the risk that the framework appears to be based only on the researcher's interpretation.

To further clarify the strength of the evidence, an evidence prominence scale was used. This scale was qualitative rather than based on exact word counts, because different papers often describe similar ideas using different terms. A source was classified as *moderate* when the concept was clearly discussed and relevant, but mainly served as supporting evidence. A source was classified as *high* when the concept was central to the paper and appeared in several important sections, such as the abstract, framework, results, or conclusion. A source was classified as *very high* when the concept formed one of the main contributions or analytical foundations of the paper and was repeatedly discussed throughout the study.

This requirement-definition phase therefore strengthens the design science process in two ways. First, it creates a clearer bridge between the problem understanding and the design of the artefact. Second, it improves transparency by showing how the final framework requirements were derived from the literature. This is especially important in this study because airport turnaround digital twins are still an emerging research area, and because no existing maturity model directly fits the specific context of airport turnaround operations.

2.1.3. C. Design and Develop Artefact

SQ2 - Which maturity dimensions, levels, and assessment logic should a digital twin maturity assessment framework for airport turnaround operations contain?

This sub-question corresponds to the *Design and Develop Artefact* phase and forms the constructive core of the research.

Making of the literature-based framework

The literature-based framework was developed as the first step of the *Design and Develop Artefact* phase. The purpose of this step was to create an initial maturity assessment framework that could later be reviewed and refined through expert input. This fits the design science logic in which an artefact is developed from the problem context and requirements, and then refined through demonstration, evaluation, and feedback (Hevner et al., 2004; Peffers et al., 2007). The output of this step is presented in Chapter 5.

The development process followed the same structure as the literature-based framework chapter. First, the requirements from SQ1 were translated into a framework structure. Second, the literature base was organised into relevant streams. Third, the maturity levels were formulated. Fourth, the framework dimensions and indicators were developed. Fifth, an initial scoring logic was drafted. These steps are explained below.

The first step was to translate the requirements from Chapter 4 into framework elements. The requirements described what the framework should assess, how it should be structured, and under which airport-specific conditions it should be applicable. However, the requirements were not copied directly into the framework as separate assessment items. Instead, they were synthesised into broader framework elements, including maturity levels, dimensions, indicators, and broader design choices. This is consistent with design science research, where requirements guide the construction of the artefact (Johannesson & Perjons, 2014), and with maturity model development, where design inputs are structured into levels, dimensions, indicators, and assessment criteria (Becker et al., 2009).

To make this translation transparent, a requirement-to-framework traceability table was created. This table shows how each requirement informed one or more dimensions, indicators, or framework-wide design elements. The purpose of this traceability step was to reduce the risk that important requirements were lost during the transition from the requirement set to the framework structure. It also makes explicit where researcher interpretation was involved in the design process.

The second step was to organise the literature base used for the framework. The framework was not developed by adopting one existing maturity model directly. Existing digital twin maturity models are often developed for other domains, such as manufacturing, asset management, infrastructure, smart cities, or general digital twin implementation. Although these models provide useful theoretical foundations, they do not fully address the specific characteristics of airport turnaround operations. Therefore, the framework was developed through synthesis and contextual adaptation.

Three streams of literature were used in this synthesis. First, general digital twin maturity literature was used to identify common maturity progressions from basic digital representation towards monitoring, prediction, prescription, optimisation, and, in some cases, autonomy (Kritzinger et al., 2018; Y. Liu et al., 2024; Madni et al., 2019). Second, detailed digital twin maturity models were used to understand how maturity can be operationalised through dimensions, indicators, rubrics, and scoring logic (L. Chen et al., 2021; Li et al., 2024; Uhlenkamp et al., 2022). Third, airport- and turnaround-related digital twin literature was used to contextualise the framework for airport turnaround operations, including operational visibility, stakeholder coordination, system integration, delay prediction, dynamic rescheduling, resource allocation, and operational decision-making (Attar et al., 2024; Conde et al., 2022; C. Liu et al., 2024; Luo et al., 2024).

The third step was to formulate the maturity levels. The reviewed maturity literature commonly describes digital twin development as a progression from a basic digital model towards more integrated, predictive, prescriptive, and autonomous capabilities. This logic was used as the foundation for the maturity-level structure. However, the maturity logic was adapted to the airport turnaround context. Several digital twin maturity models describe autonomy as a high or final maturity stage, but full autonomous control was not selected as the endpoint in this framework. Airport turnaround operations are safety-critical, time-constrained, and dependent on coordination between multiple responsible stakeholders. Therefore, the highest level in the initial framework was defined as optimisation-oriented decision support rather than full autonomous control. This means that a highly mature turnaround digital twin should support proactive, data-driven, and optimisation-based decision-making, while maintaining human oversight and organisational accountability.

The fourth step was to define the framework dimensions and indicators. The maturity levels provide the overall progression of the framework, but they do not specify which aspects of the digital twin should be assessed. Therefore, detailed maturity models were used to identify recurring capability areas, such as data, integration, analytics, decision support, operational representation, human interaction, governance, interoperability, and trust (L. Chen et al., 2021; Li et al., 2024; Uhlenkamp et al., 2022). These capability areas were then combined with the requirements from SQ1 and with airport-specific literature to create dimensions that fit the turnaround context.

During this process, several requirements had to be synthesised rather than transferred one-to-one. For example, real-time operational visibility is related to the operational representation of the turnaround process, but also depends on data availability and system integration. Similarly, decision support, resource optimisation, and dynamic rescheduling are closely related and were therefore translated into a broader dimension concerning decision support and optimisation. This synthesis resulted in a multi-dimensional framework that covers both technical and organisational aspects of digital twin maturity.

Indicators were then formulated for each dimension to make the framework assessable in practice. The indicators were based on recurring capabilities identified in the reviewed literature and on the requirements derived earlier in the research. They were written in a way that allows differences between maturity levels to be assessed. This was necessary because the framework is intended not only to describe digital twin maturity, but also to identify improvement priorities.

The fifth step was to develop an initial scoring logic. A dimension-based and indicator-based scoring approach was selected because digital twin maturity may develop unevenly across different areas. For example, an airport may have a relatively advanced data infrastructure while still having limited predictive or decision-support capability. Assessing maturity per dimension therefore provides a more detailed and actionable result than assigning only one overall maturity label. This approach is consistent with maturity model literature, where maturity models are used not only to classify a current state, but also to identify improvement areas and guide further development (Becker et al., 2009; Uhlenkamp et al., 2022).

In addition, a preliminary threshold logic was drafted to connect indicator scores to maturity-level assignment. This means that selected indicators need to reach a minimum score before a maturity level can be assigned.

The purpose of this draft scoring logic was to make the assessment more transparent and to prevent strong performance in one area from fully compensating for weak performance in an essential area. At this stage, the scoring logic was treated as an initial proposal rather than as a fully validated measurement instrument.

The output of this step was the initial literature-based maturity assessment framework presented in Chapter 5. This initial framework consists of requirement-to-framework traceability, a literature basis, maturity levels, framework dimensions, indicators, and preliminary scoring logic. It should be understood as a preliminary artefact, because it was based on literature and researcher synthesis only. In the next design step, the framework was reviewed through expert input to assess its relevance, completeness, understandability, and applicability to airport turnaround operations.

Interviews on literature-based framework
This study used qualitative expert interviews and one focus group to refine the literature-based digital twin maturity framework. A qualitative approach was suitable because the aim was not to test the model with numbers, but to understand how experts interpreted the model, which parts were unclear, and how useful the model could be in the airport turnaround context. Qualitative research is useful for studying expert views, meaning, context, and practical concerns(Patton, 2015).

Participants were selected through purposive sampling. This means that participants were chosen because they could give relevant and useful feedback on the framework (Patton, 2015). In total, eight experts participated: one researcher, two consultants, three Schiphol participants, and two KLM participants. The researcher was included to reflect on the academic structure of the model. Schiphol participants were included because the model is intended to be used in the Schiphol turnaround context. KLM participants were included because airlines are important actors in the turnaround process and because some operational data may come from airline systems. Consultant 2 was included because of broader aviation expertise in the Netherlands. Consultant 1 was included because of experience with digital twin development and digital twin consultancy.

The number of participants was not based on statistical representativeness. Instead, the aim was to collect useful expert feedback from different perspectives. In qualitative research, sample size is often linked to saturation. Saturation means that new interviews no longer add many new themes or insights (Guest et al., 2006). Guest et al. (2006) found that many main themes can appear within the first six interviews, while saturation may occur around twelve interviews in focused studies. Hennink et al. (2017) also show that fewer interviews may be enough to identify main themes, while more interviews may be needed to understand these themes in more depth. In this study, eight participants were considered suitable because the aim was to refine an existing framework, not to develop a completely new theory. The sample covered the main perspectives needed for this phase: academic structure, digital twin expertise, aviation expertise, Schiphol operations, airline involvement, and data-related issues.

The data were collected through semi-structured interviews and one focus group. Semi-structured interviews were suitable because they gave structure to the conversation, while still allowing follow-up questions and new insights (Adams, 2015). The literature-based framework was used as the interview guide. Therefore, the interviews were structured around the maturity levels, dimensions, indicators, and practical use of the framework. Participants were asked whether the model was logical, clear, complete, and usable in practice.

Three one-on-one expert interviews were conducted. The interview with the researcher was conducted online and focused mainly on the conceptual structure of the model. Consultant 1 was also interviewed online and gave feedback from a digital twin perspective. Consultant 2 was interviewed in person and gave feedback from a broader aviation perspective. One-on-one interviews were useful because each expert could explain their own view in detail.

An in-person focus group was conducted with three Schiphol participants from different parts of the operation and the digital twin context. A focus group was suitable because the model is intended to be used in an operational setting where different perspectives interact. The group setting allowed participants to respond to each other, clarify practical examples, and discuss how the model could be used at Schiphol. Focus groups are useful when interaction between participants can create extra insight (Kitzinger, 1995; Morgan, 1996).

An online paired expert interview was conducted with two KLM participants. This was not labelled as a focus group because it involved only two participants from the same organisation. A paired interview was suitable because the two participants could complement each other and provide a fuller view of the airline perspective. Paired interviews are useful when participants share a context but can add different details or examples (Morgan, 1996).

Both online and in-person formats were used. This choice was mainly practical and depended on participant availability and location. Online interviews were useful for participants who were not available in person. In-person meetings supported direct discussion of the framework. Online interviews can be a valid method for qualitative data collection when participants are comfortable with digital communication tools (Archibald et al., 2019).

The interviews were conducted in English or Dutch. The interviews with the Researcher, Consultant 1, and KLM were conducted in English. For most participants and the interviewer, English was a second language. The interview with Consultant 2 and the Schiphol focus group were conducted in Dutch, which was the participants' mother tongue. This helped participants explain their ideas clearly. Dutch input was translated into English where needed for reporting. Because translation can influence meaning in qualitative research, the translation focused on the meaning of the response rather than literal wording (Temple & Young, 2004; Van Nes et al., 2010).

The interview data were analysed thematically. The analysis combined deductive and inductive coding. First, feedback was coded deductively using the structure of the literature-based framework, such as maturity levels, dimensions, and indicators. Second, inductive codes were added when new themes emerged from the interviews, such as unclear wording, overlap between indicators, commercial sensitivity, decision rights, workflow integration, and safety boundaries. A combined deductive and inductive coding approach is suitable when research starts from an existing framework but also remains open to new themes from the data (Fereday & Muir-Cochrane, 2006).

After coding, the feedback was compared across participants. Recurring comments, contradictions, missing elements, and practical concerns were identified. Each part of the framework was then reviewed to decide whether it should be kept, clarified, renamed, merged, removed, or moved to another part of the model. This made the refinement process systematic and transparent.

Interview on refined framework

After the literature-based framework had been refined, a follow-up expert interview was conducted with the same researcher who had reviewed the first version of the framework. The researcher was selected through purposive sampling because of their expertise in digital twins and maturity assessment. In qualitative research, purposive sampling is appropriate when participants are selected because they can provide information-rich insights related to the research aim (Patton, 2015).

The follow-up interview was used as a methodological step to review the refined framework before it was applied in the next phase of the study. This fits the logic of design science research, where an artefact is developed, evaluated, and refined through an iterative process (Hevner et al., 2004; Peffers et al., 2007).

The interview was semi-structured. The refined framework was used as the interview guide, so the discussion focused on the main parts of the model. At the same time, the semi-structured format allowed follow-up questions and additional comments from the researcher. Semi-structured interviews are useful when the researcher wants to discuss the same core topics, while still leaving room for clarification and unexpected insights (Adams, 2015).

Using the same researcher for the follow-up interview created continuity in the expert review process. The researcher was already familiar with the earlier version of the framework and the feedback given in the first review round. This supported an iterative evaluation process, in which the artefact was reviewed, refined, and reviewed again before being used in the next research phase (Hevner et al., 2004; Peffers et al., 2007).

The interview data were used as expert feedback on the refined framework. The feedback was analysed in relation to the structure and assessment logic of the framework. This helped to assess whether the framework was ready to be used in the next research phase.

2.1.4. D. Demonstrate Artefact

SQ3 - To what extent can the proposed maturity assessment framework be meaningfully applied in a real airport turnaround context?

This sub-question addresses the *Demonstrate Artefact* phase. The aim of this phase was to show that the final digital twin maturity framework could be used in practice. In design science research, demonstration is used to show how an artefact works in the context for which it was developed (Peffers et al., 2007). In this study, the framework was demonstrated in the Schiphol turnaround context.

A single-case demonstration was used. Schiphol was selected because the framework was developed for airport turnaround operations and because Schiphol provided a relevant practical case. The aim was not to compare airports, but to test whether the framework could support a useful maturity assessment in one real operational setting. A single case can be valuable when it gives detailed, context-specific insight into the situation being studied (Flyvbjerg, 2006).

The demonstration was conducted as a focus group application session with five Schiphol participants. Three participants had also joined the earlier framework refinement session. They were included again because they already knew the framework and could provide continuity between the refinement and demonstration phases. Two additional participants were included to add new operational and organisational perspectives.

A focus group was suitable because the assessment required discussion between participants with different perspectives. Focus groups are useful when interaction helps participants explain, compare, and develop their views (Kitzinger, 1995; Morgan, 1996). This was important because the maturity assessment was not only based on individual judgement, but also on discussion about how the framework should be interpreted in the Schiphol context.

The final framework was used as the structure for the session. First, the external implementation constraints were discussed together. This created a shared understanding of the practical context before the internal maturity was scored. The discussion included data access, commercial sensitivity, decision rights, workflow integration, and safety boundaries.

After this, the internal maturity indicators were scored. Participants first scored the indicators individually. This allowed each participant to make their own judgement before the group discussion. Afterwards, the scores were discussed as a group. When participants gave different scores, the differences were discussed until one shared score was agreed. This approach combined individual judgement with group discussion. This is useful because individual scoring can reduce the risk that one participant immediately influences the others, while group discussion helps clarify different interpretations and reach a shared assessment. This is also consistent with the logic of the Nominal Group Technique, where individual input is followed by clarification, discussion, and rating or agreement (McMillan et al., 2016).

The final maturity scores were therefore based on both individual scoring and collective discussion. The aim was not only to produce a numerical score, but also to understand the reasoning behind the score. The discussion around different interpretations was therefore treated as an important part of the demonstration data.

After the scoring, possible development steps were identified. These were based on the difference between the current maturity scores and the requirements for higher maturity levels. Indicators with lower scores, or external constraints that limited further development, were treated as areas where improvement would be needed. In this way, the framework was used to describe the current maturity level and to discuss what would be needed to develop further.

The demonstration had two purposes. First, it tested whether the framework could be applied in a real airport turnaround setting. Second, it tested whether the framework could support a structured discussion about maturity, limitations, and future development steps. This fits design science research, where artefacts are demonstrated and evaluated in their problem context (Hevner et al., 2004; Peffers et al., 2007).

2.1.5. E. Evaluate Artefact

SQ4 - To what extent is the digital twin maturity assessment framework applicable and useful for airport turnaround operations?

This sub-question addresses the *Evaluate Artefact* phase. The aim was to assess to what extent the digital twin maturity assessment framework was applicable and useful for airport turnaround operations. In design science research, evaluation assesses whether an artefact fulfils its intended purpose and addresses the problem for which it was designed (Hevner et al., 2004; Peffers et al., 2007). This is also consistent with the DSR logic that artefacts should be evaluated in relation to the objectives and requirements that guided their development (Johannesson & Perjons, 2014; Peffers et al., 2007).

The evaluation consisted of two parts: a Schiphol evaluation and an external applicability check with Vancouver International Airport (YVR). The Schiphol evaluation was based on the transcript of the demonstration session and the completed parts of the post-demonstration questionnaire. The transcript formed the main qualitative evidence, while the questionnaire was used as supplementary evidence because it was only partly completed. This qualitative approach was suitable because the aim was to understand how participants of the demonstration interpreted the framework in context, rather than to statistically test predefined variables (Patton, 2015).

The Schiphol transcript was separated into two types of material. Score-related discussion was used to support the demonstration results in Chapter 8, because these fragments explained why certain external constraint levels and internal maturity scores were assigned. Model-related discussion was used for the evaluation in Chapter 9, because these fragments concerned the clarity, usefulness, completeness, and practical applicability of the framework itself.

The questionnaire was directly linked to the requirements defined in SQ1. It did not ask participants to reassess Schiphol's maturity, but to evaluate the framework itself, including the maturity levels, dimensions, indicators, scoring logic, and the distinction between internal maturity and external implementation constraints. Because several questionnaire answers were missing, the questionnaire was not analysed as a complete quantitative survey. Instead, the completed answers and written comments were treated as supplementary qualitative evidence.

The Schiphol evaluation data were analysed qualitatively. The model-related transcript fragments and questionnaire comments were grouped according to the main evaluation topics: clarity, usefulness, completeness, applicability, the distinction between internal maturity and external constraints, and the use of the framework for identifying future development steps. Additional themes were included when they emerged from the data. This combined deductive and inductive approach was suitable because the analysis started from the existing framework and requirements, but also remained open to new themes (Fereday & Muir-Cochrane, 2006).

In addition, an online semi-structured interview was conducted with a representative from YVR. This interview was used as an external applicability check. The aim was not to score YVR with the framework, but to explore whether the framework was understandable, relevant, and meaningful outside the Schiphol context. Semi-structured interviews are suitable when the same core topics need to be discussed while allowing room for context-specific explanation and feedback (Adams, 2015).

YVR was selected because it differs from Schiphol in the organisation of turnaround operations and in the maturity of turnaround-related digital twin capabilities. This follows the logic of purposeful and contrasting case selection, where a case is selected because it provides relevant contextual variation rather than statistical representativeness (Seawright & Gerring, 2008). Therefore, YVR was not used to replicate the Schiphol assessment, but to examine whether the framework could also be understood and discussed in another airport setting.

Because the Schiphol questionnaire was incomplete and the external evaluation was based on one interview, the evaluation should be interpreted as an exploratory qualitative artefact evaluation rather than as a statistically validated assessment.

2.2. Research Design

The methodological process described above is summarised in Figure 2.1. The figure provides an overview of the research flow by showing how the DSR phases are operationalised through the four sub-questions and how these phases are connected throughout the study.

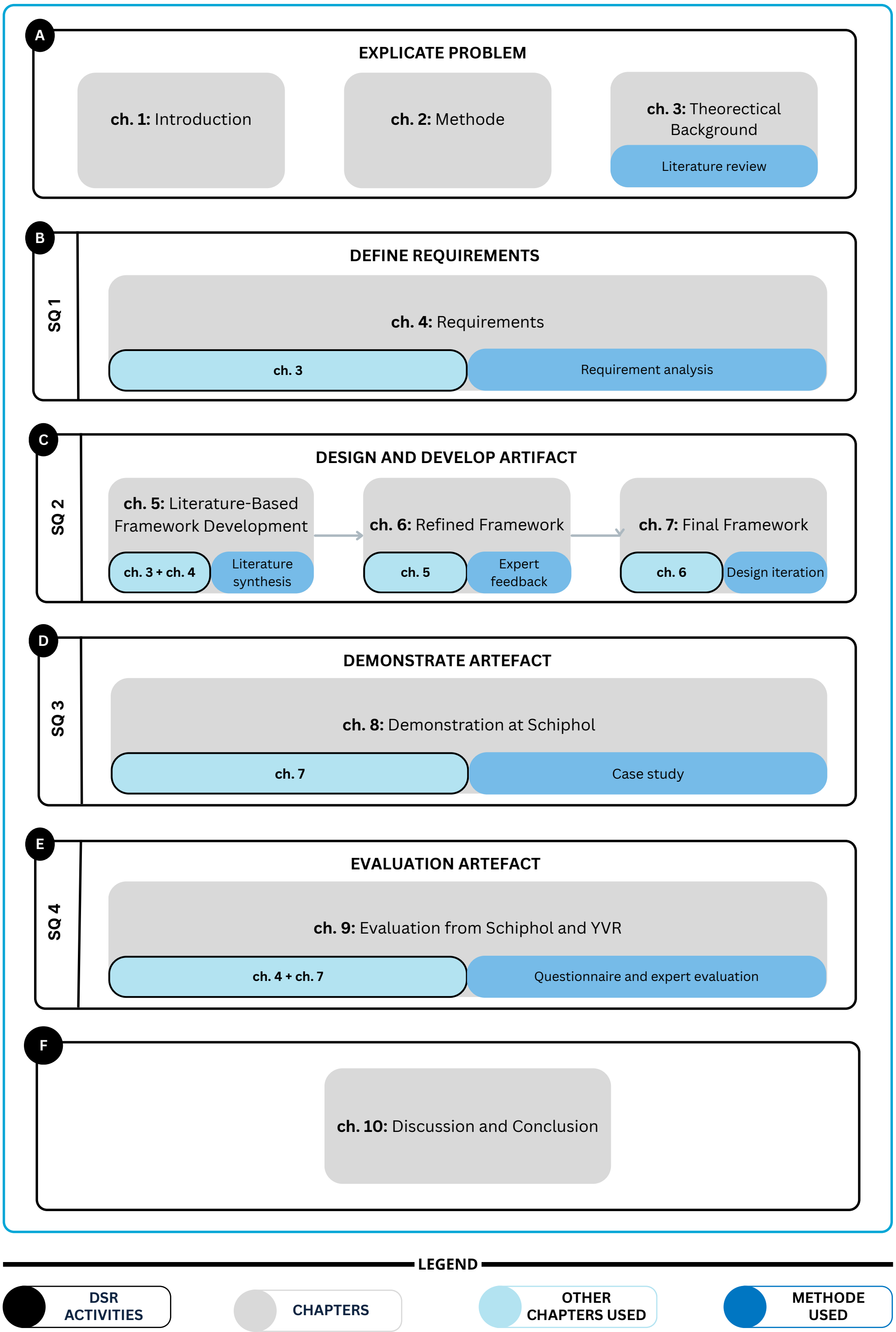


Figure 2.1: Research design and Design Science Research flow

3
Theoretical Background

This chapter provides the theoretical foundation for the development of the digital twin maturity assessment framework. It first introduces the key concepts underlying this study, including digital twins, airport turnaround operations, and digital twin maturity. Subsequently, it reviews existing literature on digital twin applications in airport contexts and maturity models from other domains. Based on this review, the chapter identifies the main limitations of existing approaches and formulates the research gap that guides the design of the proposed framework.

3.1. Background Knowledge

This section introduces the main terms and concepts used throughout this research. The purpose is not to provide a full literature review, but to create a common understanding of the operational and technical terminology on which the study builds.

3.1.1. Aircraft turnaround process

Aircraft turnaround refers to the ground process in which an arriving aircraft is prepared for its next departure. In this study, the turnaround is defined as the process from on-block to off-block. On-block refers to the moment the aircraft has arrived at its assigned stand or gate and can be secured for ground handling activities. Off-block refers to the moment the aircraft is ready to leave the stand, usually for pushback or taxi-out. Landing, taxi-in, taxi-out and take-off are therefore treated as related arrival and departure processes, but not as part of the core turnaround itself.

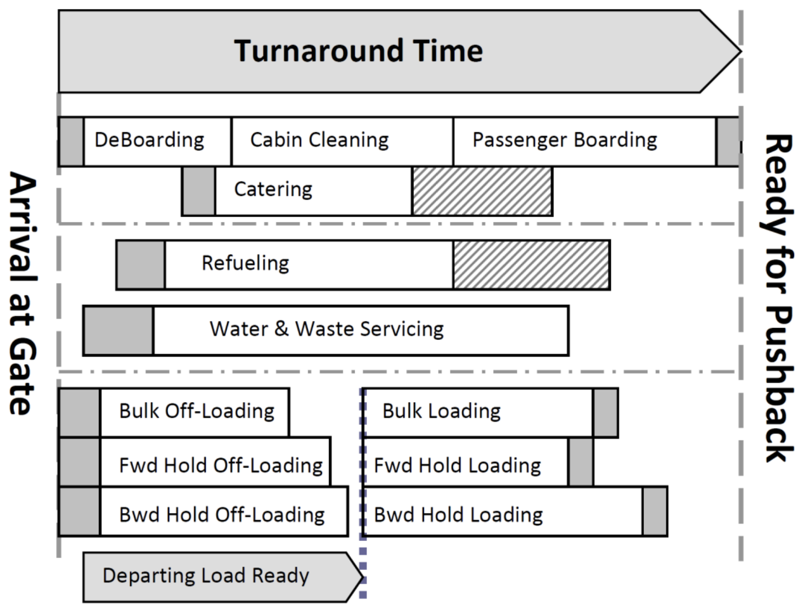


Figure 3.1: Generic aircraft turnaround process chart, adapted from Fuchte et al. (2012).

Figure 3.1 shows the turnaround as the time between arrival at the gate and readiness for pushback. During this period, several ground handling activities must be completed. The cabin-related process consists of passenger deboarding, cabin cleaning and passenger boarding. In parallel, the aircraft may be catered, refuelled, and serviced with water and waste handling. At the same time, inbound baggage or cargo is off-loaded and outbound baggage or cargo is loaded into the aircraft holds.

The figure also illustrates that turnaround is not a simple linear process. Some activities have to take place in sequence, such as deboarding before boarding, while other activities can occur partly in parallel, such as refuelling, catering, servicing and baggage handling. However, these activities remain interdependent because they take place around the same aircraft and within the same limited time window. This makes the turnaround a time-critical coordination process involving multiple actors, such as the airport operator, airline, ground handler and air traffic control (Okwir et al., 2017). Inefficiencies in the turnaround can contribute to flight delays and may propagate disruptions through the wider airline network (Evler et al., 2022; Wu & Caves, 2000). For this reason, the turnaround process forms the operational scope of this research.

3.1.2. Airport Collaborative Decision Making

Airport Collaborative Decision Making (A-CDM) is a coordination concept for managing airport operations through shared and timely operational information. In the turnaround context, A-CDM does not change the physical activities themselves, such as unloading, cleaning, fuelling, boarding, and pushback preparation. Instead, it structures how the involved actors share information about the progress of these activities. The concept links the inbound flight to the outbound flight and follows the aircraft through a sequence of operational milestones, from arrival and in-block status to aircraft readiness, start-up approval, off-block time, and take-off planning (EUROCONTROL, n.d.).

In practice, A-CDM requires airport partners, such as the airport operator, aircraft operator, ground handler, air traffic control, and Network Manager, to work with common estimates and target times. Important examples are the Target Off-Block Time (TOBT) which indicates when the aircraft is expected to be ready to leave the stand, and the Target Start-Up Approval Time (TSAT), which indicates when start-up approval is expected to be given. These times are updated when the turnaround situation changes. For example, if boarding, fuelling, cleaning, baggage loading, or de-icing takes longer than expected, the TOBT can be adjusted. This update can then influence the departure sequence, start-up planning, and wider network planning (EUROCONTROL, n.d.).

A-CDM therefore suggests a structured way of coordinating the turnaround through shared milestones, common time definitions, consistency checks, and updates to target times. Its purpose is to ensure that all relevant actors work with the same operational picture and that changes during the turnaround are communicated early enough to improve predictability and reduce avoidable waiting time.

3.1.3. Digital twin

A digital twin is a digital representation of a physical object, system, or process that is connected to its real-world counterpart through data exchange. This representation does not have to be a complex three-dimensional model. In practice, a digital twin can also take the form of a dashboard, data platform, simulation environment, or even a structured spreadsheet, as long as it is connected to relevant operational data and represents the state or behaviour of the physical system. In contrast to a static digital model, a digital twin is updated with operational data and can therefore reflect the current state of the physical system. Depending on its maturity, a digital twin can support monitoring, simulation, prediction, optimisation, and decision support (Iliuță et al., 2024; Singh et al., 2021; Van Der Horn & Mahadevan, 2021).

In the context of airport turnaround operations, a digital twin can be understood as a data-driven representation of the turnaround process. It may help stakeholders monitor task progress, identify delays, predict operational consequences, and support decisions about resources, schedules, or interventions. In this research, the digital twin is therefore considered as a potential next step beyond information sharing, because it can support a more dynamic and predictive understanding of the turnaround process.

3.1.4. Digital twin maturity

Digital twin maturity refers to the degree to which a digital twin has developed from a basic digital representation into a more advanced system with stronger operational capabilities. At lower maturity levels, a digital twin may mainly visualise or document the physical system. At higher maturity levels, it may integrate real-time data, support monitoring and prediction, and provide optimisation-oriented decision support (L. Chen et al., 2021; Iliuță et al., 2024; Uhlenkamp et al., 2022).

A maturity assessment framework is a structured method for assessing this development. It uses maturity levels, dimensions, and indicators to determine the current state of a digital twin and to identify possible next development steps. In this research, such a framework is developed specifically for digital twins in airport turnaround operations.

3.1.5. Deep Turnaround

Deep Turnaround is a turnaround intelligence solution that uses AI-based image processing to detect and report events during the aircraft turnaround process. Instead of relying only on manually entered milestone data, it can automatically identify operational events from visual information, such as the presence or movement of ground handling equipment and the progress of turnaround activities. Royal Schiphol Group describes the solution as being able to detect more than 70 turnaround events across 30 turnaround processes and to provide historic, real-time, and predictive insights into turnaround performance (Royal Schiphol Group, 2024).

In this research, Deep Turnaround is relevant because it illustrates how detailed turnaround data can be captured and transformed into operational insight. It is therefore not treated as a complete digital twin by itself, but as a technology that can support important digital twin capabilities, especially automatic data capture, current situation visibility, historical analysis, and prediction in the turnaround context.

3.2. Literature review

This section presents the literature review that forms the theoretical foundation of this study. In line with the DSR approach, the review serves as part of the knowledge base on which the artefact is developed. To identify relevant studies in a structured and transparent manner, a systematic literature search was conducted using the Scopus database. The search and selection procedure followed a PRISMA-inspired process, so that the identification, screening, eligibility assessment, and final inclusion of studies could be reported explicitly (Page et al., 2021). The corresponding review flows are presented in Appendix A.

The aim of this review is to identify what is already known and what remains unexplored with regard to digital twins in airport turnaround operations. Because the literature on this specific topic is still limited, the review first examines the small number of digital twin turnaround-focused studies. It then broadens the perspective to airport digital twin research more generally and finally to digital twin maturity models from other domains. In doing so, the section synthesises key concepts, identifies major research themes, and establishes the academic gap that motivates this study.

3.2.1. What Slightly Exists in Literature: Digital Twin and Turnaround

The literature on digital twins specifically for aircraft turnaround operations is still very limited. Figure A.1 presents the review workflow, and table A.1 provides an overview of the selected studies. In this review, only two directly relevant peer-reviewed studies were identified, which indicates that this remains an emerging rather than a mature research field. The first study by Conde et al. (2022) develops a digital twin architecture for managing turnaround information at Aberdeen International Airport, with the aim of improving coordination, data standardisation, and real-time visibility of turnaround events to reduce delays. Its contribution lies mainly in improving information integration and shared operational awareness in a real airport setting. The second study by Lu et al. (2025) focuses on automated aircraft turnaround and uses a digital twin combined with a sandbox-based apron to simulate smart-device coordination. In this way, it shows how a digital twinh can also be used to test and improve future automated turnaround processes in an experimental environment.

What these two papers have in common is that both treat turnaround as a highly time-critical airport process in which better coordination, real-time data exchange, and stronger operational visibility can improve efficiency. At the same time, they differ in emphasis. Conde et al. (2022) focus on information management and coordination in a real airport context, whereas Lu et al. (2025) focus on automation and simulation in an experimental setting. A key difference also lies in how the two papers understand the digital twin itself. Conde et al. (2022) adopt a more conceptual interpretation, emphasizing the digital twin as a connected representation of a real operational system in which integration and data exchange are central. By contrast, Lu et al. (2025) take a more functional and engineering-oriented view, treating the digital twin primarily as a simulation and testing environment for automated turnaround operations. This distinction is relevant because it shows that turnaround-related digital twin research currently develops along two lines: one aimed at improving operational coordination in practice, and another aimed at exploring future automated turnaround concepts under controlled conditions.

Although these studies confirm the relevance of digital twins for turnaround operations, they also show how limited the turnaround-specific academic evidence base still is. This is notable because, as seen in practice and in grey literature, more airports are investing in digital twin-related initiatives and turnaround intelligence tools. However, this development is not yet matched by a strong academic literature on digital twins for this specific part of airport operations. To examine whether this gap also exists in other parts of the airport, it is therefore necessary to broaden the perspective to airport digital twin research more generally.

3.2.2. What Exists in Literature: Airport Digital Twin Research

To understand how digital twin technology is discussed in airport research, a literature review was conducted on studies combining digital twins and airports. Figure A.2 presents the review workflow, and table A.2 provides a structured overview of the selected studies. Compared with the very limited literature on turnaround-specific applications, this broader body of work shows that airport digital twins are discussed across a much wider range of domains and purposes. Within this literature, Conde et al. (2022) remain particularly relevant because they connect airport digital twin development directly to turnaround operations, where coordination and information sharing are central. At the same time, the wider airport literature shows that digital twins are also being explored for operational optimisation, simulation, decision support, human–system interaction, infrastructure and facility management, sustainability planning, and even coordination across multiple airports.

A first group of studies focuses on the operational value of digital twins for managing and improving airport processes. Luo et al. (2024), for example, move beyond information visibility by using a digital twin-based scheduling framework to support dynamic rescheduling of specialised vehicles under disruption. Their study shows how a digital twin can detect disruptions, update the virtual model with real-time data, and generate revised schedules to reduce delays and operational conflicts. Qiu (2024) similarly position the digital twin as a mechanism for dynamic decision-making, using real-time simulation to optimise resource allocation in air terminal yards. Attar et al. (2024) broaden this operational perspective to the airside as a whole, showing that a digital twin can serve as a real-time decision-support environment for analysing delay and capacity effects under different scenarios. Saifutdinov et al. (2022) add an important testing perspective by showing that digital twins can also function as an emulation environment for centralised ground traffic control, allowing airport processes and control software to be tested under realistic but still simulated conditions. Finally, Xu et al. (2025) extend operational thinking beyond the single airport by showing how a digital twin can support real-time synchronisation and optimisation of flight departures across a multi-airport region.

Taken together, these studies suggest that airport digital twin development may progress from supporting operational visibility and coordination toward simulation, rescheduling, optimisation, and increasingly system-wide forms of control. They also show that the object of the digitak twin can differ substantially, ranging from turnaround events and specialised vehicle operations to terminal yards, airside capacity, and regional airport networks. This is important because it indicates that airport digital twin development is not tied to one single application, but may evolve across different operational layers of the airport system. However, while these studies demonstrate a broadening range of capabilities, they do not translate these developments into a maturity logic that helps airports understand how such capabilities relate to one another over time.

A second group of studies shows that airport digital twin development is not only an operational issue, but also depends on human, organisational, and infrastructural conditions. K. Chen et al. (2024) demonstrate that digital twin applications in airport operations also involve questions of usability, workload, trust, and shared situational awareness, as their mixed-reality digital tower was evaluated from a human-factors perspective. Abdelalim et al. (2025) show that digital twins can support more proactive and predictive facility management by integrating real-time data into maintenance and asset management processes. Moon et al. (2025) make this organisational and infrastructural dimension even more explicit by identifying major barriers for airport facility digitalisation, including fragmented maintenance environments, the absence of data standards and systems, and insufficient user-oriented platforms. Based on these barriers, they propose a phased roadmap of development, integration, and advancement. The importance of implementation conditions is also visible in studies related to Istanbul Airport. Mizrak et al. (2024) position digital twins within a broader strategic roadmap for sustainable and digital airport management. Durmaz and Göçmen (2025) provide a more practice-oriented case study of Istanbul Airport, showing how digital twins can support infrastructure monitoring, maintenance planning, energy management, passenger-flow monitoring, operational simulation, and decision support. At the same time, their study shows that implementation depends on more than the technology itself. Issues such as data compatibility, cybersecurity, staff training, user acceptance, and coordination between departments and external stakeholders influence how far these systems can be implemented in practice. Together, these studies show that airport digital twin development depends on data infrastructure, organisational readiness, implementation strategy, and user adoption.

These studies are highly relevant because they show that airport digital twin development may involve not only turnaround coordination and operational optimisation, but also simulation, human interaction, facility management, infrastructure readiness, strategic planning, and inter-airport integration. Conde et al. (2022) still provide the most direct foundation for this thesis because they locate the digital twin in the turnaround domain itself, but the broader airport literature shows that airport digital twin development is a much wider socio-technical process. What remains missing across these studies is not evidence that airports are developing digital twin capabilities, but a structured way to evaluate and guide that development across stages. This suggests that airport digital twin development should be understood as a progression of capabilities rather than as a fixed yes-or-no state. To examine how such staged development has already been conceptualised elsewhere, the following section reviews existing digital twin maturity models from other domains.

3.2.3. What Exists in Literature: Digital Twin Maturity Models

This section reviews how digital twin maturity is conceptualized in the broader literature. Figure A.3 presents the workflow for selecting digital twin maturity model studies, and table A.3 summarizes the final set. Across this literature, maturity models are used to describe how digital twins develop over time and how their capabilities can be assessed. At the same time, the reviewed studies show that no single model can be transferred directly to airport turnaround operations, because most are developed for other domains such as smart cities, buildings, infrastructure, manufacturing, or asset management (L. Chen et al., 2021; Z.-S. Chen et al., 2024; Haraguchi et al., 2024; Li et al., 2024; Mo et al., 2023) .

A first common theme in this literature is that digital twin maturity is usually described as a staged progression. This staged view is already visible in Madni et al. (2019), who distinguish between a pre-digital twin, a digital twin, an adaptive digital twin, and an intelligent digital twin. Their model shows that maturity is not only about having a virtual model, but also about how strongly this model is connected to the physical system, how often it is updated, and whether it can support operational decision-making. This also aligns with Van Der Horn and Mahadevan (2021), who stress that a digital twin should represent a specific physical system and be updated through information exchange with that system. Hartmann and Wimmer (2026), in their comparison of smart city digital twin maturity models, show that recent models increasingly follow a similar structure, often with five or six levels. Y. Liu et al. (2024) also propose a five-level digital twin maturity model and link each level to specific capabilities such as virtual representation, two-way communication, simulation, control, lifecycle support, and intelligence. Li et al. (2024) describe a comparable five-level hierarchy for underground infrastructure, while Scheibmeir and Malaiya (2022) present a six-level model that develops from an initial digital twin to a ubiquitous digital twin that also includes wider environmental context. Haraguchi et al. (2024) extend this further in the CITYSTEPS model, which reaches real-time synchronization and autonomous implementation. Taken together, these studies suggest that digital twin maturity is commonly understood as a development from basic representation toward more integrated, predictive, adaptive, and autonomous systems.

At the same time, the literature shows that digital twin maturity is not only a matter of technical connection. Several studies argue that maturity should be assessed across multiple dimensions. Uhlenkamp et al. (2022), for example, identify 31 characteristics across seven categories, including context, data, computing capabilities, model, integration, control, and human–machine interaction. L. Chen et al. (2021) evaluate digital twin maturity in asset management using the Gemini Principles of purpose, trust, and function, which emphasize value creation, interoperability, security, and data quality. Z.-S. Chen et al. (2024) add a broader construction perspective by including organizational and lifecycle-related dimensions in their maturity model. Hartmann and Wimmer (2026) also note that full lifecycle capability is still missing in most smart city digital twin maturity models. Together, these studies show that digital twin maturity should be understood as a combination of technical, organizational, and operational capabilities rather than as a purely technical issue.

Another theme in the literature is the role of the user. Many maturity models focus mainly on data, models, and system functions, but Manickam et al. (2023) argue that digital twin maturity should also be assessed from a user perspective. Their UX-DTMM (user experiance - digital twin maturity model) is built around five pillars: understandability, usability, convenience, dependability, and delight. In their view, a digital twin is only truly mature if users can understand it, interact with it effectively, and use it for decision-making. This adds an important perspective to the maturity discussion and complements the human–machine interaction dimension identified by Uhlenkamp et al. (2022).

The literature also suggests that more advanced digital twin maturity is linked to adaptability, autonomy, and resilience. Mo et al. (2023), in their study of autonomous manufacturing systems, show that higher maturity includes capabilities such as self-configuration, self-healing, self-optimization, and self-protection. Li et al. (2024) make a similar point from an infrastructure perspective by emphasizing sensing timeliness, communication stability, and the closure of the sensing-control loop. Gourisetti et al. (2025) add that as digital twins become more connected and influential, cybersecurity, digital privacy, and trust boundaries become more important as well. This means that mature digital twins are not only more capable, but also more secure and more robust.

Finally, several studies move beyond broad maturity descriptions and propose more systematic ways to assess digital twin development. L. Chen et al. (2021) define explicit rubrics for purpose, trust, and function. Li et al. (2024) use a D-ANP approach to assign weights to different maturity criteria, and Z.-S. Chen et al. (2024) develop a structured decision-support method to combine expert judgments. These studies are important because they show that digital twin maturity assessment is increasingly becoming more transparent, quantitative, and methodologically consistent.

Despite these contributions, an important gap remains. The reviewed studies provide useful insights into staged development, multidimensional assessment, human use, autonomy, security, and evaluation methods, but they are not designed for the specific conditions of airport turnaround operations. Airport turnarounds are short-cycle, high-frequency, and safety-critical processes that depend on close coordination between multiple organizations. Existing maturity models therefore offer useful building blocks, but not a directly applicable framework for this setting. The literature thus supports the need for a domain-specific maturity assessment framework for digital twins in airport turnaround operations.

3.2.4. Synthesis of the Literature and Research Gap

Taken together, the reviewed literature leads to three main insights. First, the small number of turnaround-specific digital twin studies shows that digital twins can improve coordination, visibility, and operational optimisation in turnaround processes. Second, the broader airport literature shows that digital twin development in airports already covers several application areas, such as simulation, dynamic scheduling, facility management, and human–system interaction. Third, the maturity model literature shows that digital twin development can be understood as a staged and multidimensional process that involves technical, organisational, operational, and human aspects.

At the same time, these literature streams are not yet sufficiently connected. The turnaround literature mainly focuses on specific applications, the broader airport literature mainly discusses separate use cases or technical solutions, and existing maturity models were developed for other domains such as manufacturing, smart cities, buildings, and infrastructure. As a result, the literature does not yet offer a framework that is specifically suited to assessing and guiding digital twin development in airport turnaround operations.

This is especially important because airport turnaround operations have a number of specific characteristics. They are short-cycle, safety-critical, and highly dependent on coordination between multiple actors. These characteristics make it difficult to directly apply existing maturity models from other domains.

The literature therefore not only reveals a gap, but also underlines the relevance of this research. It shows that digital twins are increasingly important in airport operations, while a domain-specific maturity assessment framework for turnaround operations is still missing. This provides a clear academic motivation next to the practical one for the present study that leads to the central research question.

Part B

Define Requirements

4
Requirements

This chapter addresses the following sub-question:

SQ1 - What requirements should a digital twin maturity assessment framework fullfil to support the development of a digital twin for airport turnaround operations?

This chapter defines the requirements for the digital twin maturity assessment framework. In this study, a requirement is understood as a condition that the framework should fulfil in order to function as a useful maturity assessment model. The requirements do not describe the technical design of the digital twin itself. Instead, they describe what the framework should assess, how it should be structured, and under which conditions it should be applicable.

The methodological process used to derive the requirements is explained in Chapter 2. In short, the requirements were derived from the literature reviewed in Chapter 3. Relevant statements, concepts, findings, and design choices were selected from the literature and grouped into requirement themes. To make this process transparent, traceability matrices and evidence tables are used throughout this chapter.

The requirements are divided into three categories: functional, structural, and environmental requirements. This distinction follows Johannesson and Perjons (2014), who distinguish between requirements related to the functions of an artefact, its structural qualities, and its environmental qualities. In this thesis, functional requirements describe what the framework should assess in a turnaround-focused digital twin. Structural requirements describe how the framework itself should be organised. Environmental requirements describe the broader conditions under which the framework should be applicable.

The chapter first explains how the literature was used to derive the requirements. It then presents the functional requirements, followed by the structural and environmental requirements. The chapter ends with a consolidated overview of the final requirement set.

4.1. Use of Literature for Requirement Derivation

Not every source reviewed in Chapter 3 was used as direct evidence for the requirements. The literature review had a broader purpose: it was used to map the state of knowledge on airport digital twins, turnaround-related digital twins, and digital twin maturity models, and to identify the research gap. The requirements chapter used this literature more selectively. Only sources that provided clear and direct evidence for what the maturity framework should assess, how it should be structured, or under which conditions it should be applicable were included in the traceability matrices and evidence tables.

As explained in Chapter 2, an evidence prominence scale was used to indicate how strongly each source supports a specific requirement. Evidence prominence does not refer to the overall quality of the paper, but to the relevance, directness, and centrality of the source for the requirement in question. The level was assigned per requirement, meaning that the same paper could have high prominence for one requirement but only moderate prominence for another.

The evidence prominence scale was applied qualitatively and was assigned per requirement. First, each source was checked for whether the concept was explicitly discussed, how central the concept was in the paper, and whether the paper provided direct support for deriving the requirement. A source was classified as *moderate* when the concept was clearly mentioned and relevant, but mainly served as supporting or contextual evidence.

A source was classified as *high* when the concept was directly addressed in an important part of the paper, such as the framework, method, results, or discussion, and provided clear support for the requirement. A source was classified as *very high* when the concept formed one of the main contributions or analytical foundations of the paper and was discussed repeatedly throughout the study. Sources with *low* prominence were excluded from the detailed evidence tables, because they only mentioned the concept briefly, indirectly, or too generally to support the derivation of a requirement.

Some reviewed studies were therefore not used to establish requirements directly. For example, Abdelalim et al. (2025) and Durmaz and Göçmen (2025) are relevant for understanding the broader development of airport digital twins, especially in relation to facility management, infrastructure monitoring, and airport-wide digitalisation. However, their contribution is less directly connected to the specific assessment of digital twin maturity in aircraft turnaround operations. Similarly, Xu et al. (2025) is relevant for showing how digital twins can support coordination across multiple airports, but its unit of analysis is broader than the aircraft turnaround process. These studies therefore helped to contextualise the research problem and support the research gap, but they did not provide sufficiently direct or specific evidence for deriving individual requirements.

This selective use of literature is consistent with the purpose of this chapter. The aim was not to use every reviewed paper equally, but to derive requirements from the sources with the strongest and most direct relevance to the framework. The evidence prominence column therefore supports traceability by showing why some sources were used as central evidence for a requirement, while others were used only as contextual background. In this way, the requirement derivation remains transparent without suggesting that all reviewed sources contribute equally to the final requirement set.

4.2. Functional Requirements: The What

Functional requirements describe the operational capabilities that the maturity assessment framework should assess in a turnaround-focused digital twin. These requirements focus on what the digital twin should be able to support in airport turnaround operations. In addition to the academic literature, (Digital Twin Consortium, 2024) was included as grey literature because it provides practical support for the airport-wide and system-of-systems perspective behind several functional requirements.

Table 4.1 provides a compact overview of the literature supporting each functional requirement. The detailed evidence for each requirement is presented in the separate tables below.

Table 4.1: Traceability matrix linking functional requirements to supporting literature

Requirement	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	P15	P16	P17	P18	P19	P20	P21	P22	P23	P24
RF1			X	X	X	X																		
RF2	X	X	X				X																	
RF3				X		X		X																
RF4				X	X																X		X	
RF5								X		X						X								
RF6								X		X														
RF7											X												X	

Requirement codes: RF1 = Real-time operational visibility; RF2 = Integrated and semantically consistent data foundation; RF3 = Proactive operational intelligence; RF4 = Decision support for turnaround interventions; RF5 = Resource optimisation; RF6 = Dynamic rescheduling capability; RF7 = Shared situational awareness.

Paper legend: P1 = (L. Chen et al., 2021); P2 = (Moon et al., 2025); P3 = (Conde et al., 2022); P4 = (Attar et al., 2024); P5 = (Van Der Horn & Mahadevan, 2021); P6 = (C. Liu et al., 2024); P7 = (Uhlenkamp et al., 2022); P8 = (Luo et al., 2024); P9 = (Manickam et al., 2023); P10 = (Qiu, 2024); P11 = (K. Chen et al., 2024); P12 = (Haraguchi et al., 2024); P13 = (Z.-S. Chen et al., 2024); P14 = (Li et al., 2024); P15 = (Gourisetti et al., 2025); P16 = (Lu et al., 2025); P17 = (Mo et al., 2023); P18 = (Scheibmeir & Malaiya, 2022); P19 = (Mizrak et al., 2024); P20 = (Hartmann & Wimmer, 2026); P21 = (Saifutdinov et al., 2022); P22 = (Y. Liu et al., 2024); P23 = (Digital Twin Consortium, 2024); P24 = (Madni et al., 2019).

4.2.1. RF1: Real-time operational visibility

A first requirement is real-time operational visibility (RF1). In the turnaround context, the framework should assess whether the digital twin provides a current and continuously updated view of aircraft status, resource use, and operational disruptions. This includes information such as whether boarding has started, whether fuelling is delayed, or whether ground equipment is available at the stand. The evidence supporting this requirement is presented in table 4.2.

Table 4.2: Literature evidence for RF1: real-time operational visibility

Code	Key evidence / citation from source	Relevant paper	Evidence prominence
RF1	Digital twins are defined as virtual representations of physical systems that are updated through the exchange of information between physical and virtual systems, which provides the conceptual basis for real-time operational visibility.	(Van Der Horn & Mahadevan, 2021)	Very high
RF1	The digital twin maps real-time sensing and business data to reflect the current operational state of the airport and enables monitoring of dynamic flight ground service processes.	(C. Liu et al., 2024)	High
RF1	The implementation includes an application that shows the real state of the airport, combining 2D and 3D representations to help ground operators monitor and schedule flights in real time.	(Conde et al., 2022)	High
RF1	Digital twins provide real-time demonstration of system performance, allowing for the continuous monitoring of throughput and aircraft waiting times on the airside.	(Attar et al., 2024)	Moderate

4.2.2. RF2: Integrated and semantically consistent data foundation

A second requirement is an integrated and semantically consistent data foundation (RF2). Turnaround operations involve many actors and systems, including airlines, ground handlers, fuelling services, and airport IT systems. These actors generate data in different formats and often use different operational definitions. The framework should therefore assess whether the digital twin combines these sources into a coherent, timely, and consistently defined representation of operations. Table 4.3 provides the literature evidence used to support this requirement.

Table 4.3: Literature evidence for RF2: integrated and semantically consistent data foundation

Code	Key evidence / citation from source	Relevant paper	Evidence prominence
RF2	Data integration technology serves as a cornerstone, integrating facility lifecycle data through Common Data Environments (CDE) and OpenBIM to ensure interoperability across stakeholders.	(Moon et al., 2025)	Very high
RF2	A central challenge is to define architectures that allow the management and homogenization of heterogeneous data collected from various airport actors using standards like FIWARE and NGSI-LD.	(Conde et al., 2022)	High
RF2	Integration is defined as one of the core maturity categories of digital twins, assessing how different entities and source information systems are connected and coordinated.	(Uhlenkamp et al., 2022)	High
RF2	A digital twin maturity model for asset management emphasizes the integration of heterogeneous digital products and asset data into an integrated and semantically meaningful structure.	(L. Chen et al., 2021)	Moderate

4.2.3. RF3: Proactive operational intelligence

A further requirement is proactive operational intelligence (RF3). Beyond showing the current state, the framework should assess whether the digital twin can identify emerging turnaround problems, diagnose likely causes, and estimate their operational consequences. For example, if fuelling is delayed, the digital twin should be able to indicate whether this is likely to affect departure time or create conflicts with later operations. The literature basis for this requirement is shown in table 4.4.

Table 4.4: Literature evidence for RF3: proactive operational intelligence

Code	Key evidence / citation from source	Relevant paper	Evidence prominence
RF3	The system enables real-time diagnosis of delays and uses a dynamic Bayesian model to analyze propagation paths and provide early warning information to personnel.	(C. Liu et al., 2024)	Very high
RF3	Simulation-based twins can predict the future outcomes of managerial decisions and support early identification of operational bottlenecks and their likely consequences.	(Attar et al., 2024)	High
RF3	By identifying anomalies such as vehicle resource shortages or equipment failures, the twin issues proactive alerts and analyzes the consequences before they impact the turnaround.	(Luo et al., 2024)	Moderate

4.2.4. RF4: Decision support for turnaround interventions

Closely related to proactive operational intelligence is decision support for turnaround interventions (RF4). The framework should assess whether the digital twin can support the selection and evaluation of appropriate operational responses when disruptions occur. For example, if a delay is detected, the digital twin could support decisions such as reallocating staff, adjusting the turnaround sequence, or prioritising certain activities. The sources that informed this requirement are listed in table 4.5.

Table 4.5: Literature evidence for RF4: decision support for turnaround interventions

Code	Key evidence / citation from source	Relevant paper	Evidence prominence
RF4	A digital twin system of systems supports airport-wide decision intelligence by linking multiple airport subsystems and enabling coordinated interventions across the airport ecosystem.	(Digital Twin Consortium, 2024)	Very high
RF4	The twin serves as a decision support tool for airport traffic control by generating instructions for resolving "critical situations" through automated algorithms.	(Saifutdinov et al., 2022)	High
RF4	Digital twin visualization supports decision-makers by improving interpretation of system data and increasing confidence in taking timely operational actions.	(Van Der Horn & Mahadevan, 2021)	High
RF4	Simulation-based twins allow managers to predict future outcomes of decisions via "what-if" scenarios, leading to an 18% reduction in expected delays.	(Attar et al., 2024)	Moderate

4.2.5. RF5: Resource optimisation

Another requirement is resource optimisation (FR5). Turnaround operations are highly time-sensitive and depend on the availability of staff, vehicles, and stand-related equipment. The framework should therefore assess whether the digital twin can support the efficient allocation and use of these resources. For example, the digital twin should help decision-makers identify underused or overloaded resources and improve their deployment across turnaround activities. Table 4.6 shows how this requirement is supported by the reviewed literature.

Table 4.6: Literature evidence for RF5: resource optimisation

Code	Key evidence / citation from source	Relevant paper	Evidence prominence
RF5	A multi-type vehicle collaborative scheduling model is developed to optimize specialized airport vehicle allocation by minimizing vehicle travel distance and waiting time in airport ground operations.	(Luo et al., 2024)	Very high
RF5	Employs mixed-integer programming and adaptive algorithms to optimize yard resource allocation, specifically minimizing transport distances and operational costs.	(Qiu, 2024)	High
RF5	Uses a network planning technique to establish coordinated operation rules among smart devices, predicting exact enter and exit timings to reduce turnaround duration by 24%.	(Lu et al., 2025)	Moderate

4.2.6. RF6: Dynamic rescheduling capability

Related to resource optimisation is dynamic rescheduling capability (RF6). Because turnaround operations are frequently affected by delays and changing conditions, the framework should assess whether the digital twin can support timely reallocation and rescheduling of activities and resources. For example, if a ground handling team is delayed at one aircraft, the digital twin should support assigning another team or adjusting the schedule of service vehicles in response. The relevant evidence for this requirement is provided in table 4.7.

Table 4.7: Literature evidence for RF6: dynamic rescheduling capability

Code	Key evidence / citation from source	Relevant paper	Evidence prominence
RF6	When deviations occur in real operational states, the twin captures dynamic airport data and triggers an event-driven rescheduling strategy to continuously adjust schedules and rapidly generate optimal solutions for multi-type vehicle coordination in response to flight delays, vehicle malfunctions, and other unforeseen operational events.	(Luo et al., 2024)	Very high
RF6	Digital twins optimize yard resources through iterative correction and linkage decision-making mechanisms that react instantly to dynamic operational factors.	(Qiu, 2024)	High

4.2.7. RF7: Shared situational awareness

A final functional requirement is shared situational awareness (RF7). Turnaround operations involve many actors who perform tightly connected tasks under time pressure. The framework should therefore assess whether the digital twin enables relevant stakeholders to work from a common and up-to-date operational picture. For example, airlines, ground handlers, and airport operators should be able to access the same status information and build a shared understanding of the operational situation. The literature support for this final functional requirement is presented in table 4.8.

Table 4.8: Literature evidence for RF7: shared situational awareness

Code	Key evidence / citation from source	Relevant paper	Evidence prominence
RF7	A unified “Airport-as-a-System” framework promotes a team-oriented approach by allowing stakeholders to visualize and act upon data-driven insights together.	(Digital Twin Consortium, 2024)	Very high
RF7	Tangible and mixed-reality interfaces with shared visualizations improve shared situational awareness and trust among air traffic controllers by enabling multiple controllers to work from the same operational picture and support collaborative decisions without excessive verbal coordination.	(K. Chen et al., 2024)	High

4.3. Structural Requirements: The Inner Workings

Structural requirements describe how the maturity assessment framework itself should be organised. They concern the internal logic of the framework, including how maturity is divided into levels, how different assessment dimensions are separated, and how scores can be assigned in a consistent way. Together, these structural requirements ensure that the framework can be used not only to describe digital twin maturity, but also to assess, compare, and guide its further development.

Table 4.9 provides a compact overview of the literature supporting each structural requirement. The detailed evidence for each requirement is presented in the separate tables below.

Table 4.9: Traceability matrix linking structural requirements to supporting literature

Requirement	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	P15	P16	P17	P18	P19	P20	P21	P22	P23	P24
RS1	X											X										X		X
RS2							X						X											
RS3							X							X										
RS4	X													X										
RS5	X												X				X							
RS6	X	X																						

Requirement codes: RS1 = Multi-level maturity structure; RS2 = Multi-dimensional assessment structure; RS3 = Modularity; RS4 = Operationalisability and scoring clarity; RS5 = Balance between technical and organisational maturity; RS6 = Actionable and prioritised progression pathway.

Paper legend: P1 = (L. Chen et al., 2021); P2 = (Moon et al., 2025); P3 = (Conde et al., 2022); P4 = (Attar et al., 2024); P5 = (Van Der Horn & Mahadevan, 2021); P6 = (C. Liu et al., 2024); P7 = (Uhlenkamp et al., 2022); P8 = (Luo et al., 2024); P9 = (Manickam et al., 2023); P10 = (Qiu, 2024); P11 = (K. Chen et al., 2024); P12 = (Haraguchi et al., 2024); P13 = (Z.-S. Chen et al., 2024); P14 = (Li et al., 2024); P15 = (Gourisetti et al., 2025); P16 = (Lu et al., 2025); P17 = (Mo et al., 2023); P18 = (Scheibmeir & Malaiya, 2022); P19 = (Mizrak et al., 2024); P20 = (Hartmann & Wimmer, 2026); P21 = (Saifutdinov et al., 2022); P22 = (Y. Liu et al., 2024); P23 = (Digital Twin Consortium, 2024); P24 = (Madni et al., 2019).

4.3.1. RS1: Multi-level maturity structure

An important structural requirement is a multi-level maturity structure (RS1). The framework should contain distinct maturity levels that represent meaningful progression in digital twin capability. Lower levels may represent basic data visualisation, while higher levels may include predictive capabilities, intervention support, and more advanced operational integration. The evidence base for this staged structure is shown in table 4.10.

Table 4.10: Literature evidence for RS1: multi-level maturity structure

Code	Key evidence / citation from source	Relevant paper	Evidence prominence
RS1	The study proposes an eight-stage maturity model (CITYSTEPS), ranging from Conception & Planning (Stage 1) to Proactive Real-time Synchronization and Self-actualized Implementation (Stage 8).	(Haraguchi et al., 2024)	Very high
RS1	This paper introduces a five-level Digital Twin Maturity Model (DTMM), which systematically aligns digital twin capabilities, phased objectives, and technical requirements within a unified framework.	(Y. Liu et al., 2024)	High
RS1	The model defines four levels of a virtual representation (Pre-Digital Twin, Digital Twin, Adaptive Digital Twin, and Intelligent Digital Twin), each with a specific purpose and scope to help with decision-making.	(Madni et al., 2019)	High
RS1	The model defines five distinct maturity stages—Unaware, Identifiable, Aware, Communicative, and Interactive—based on the ICE asset information maturity scales, offering an explicit staged structure for digital twin development.	(L. Chen et al., 2021)	Moderate

4.3.2. RS2: Multi-dimensional assessment structure

A related requirement is a multi-dimensional assessment structure (RS2). Maturity should not be reduced to a single overall score, because digital twin development often progresses unevenly across different aspects such as data integration, analytics, and user interaction. For instance, an airport may have strong data integration but still have limited decision support capabilities. Table 4.11 presents the literature supporting this multi-dimensional assessment logic.

Table 4.11: Literature evidence for RS2: multi-dimensional assessment structure

Code	Key evidence / citation from source	Relevant paper	Evidence prominence
RS2	The framework assesses Digital Twins in seven categories (context, data, computing capabilities, model, integration, control, human-machine interface) with 31 ranked characteristics.	(Uhlenkamp et al., 2022)	Very high
RS2	A BDT maturity model including 7 dimensions and 22 indicators is proposed, allowing for a comprehensive measurement rather than a single overall score.	(Z.-S. Chen et al., 2024)	Very high

4.3.3. RS3: Modularity

The framework should also be modular (RS3). This means that different digital twin capabilities or components should be assessable separately, so that specific strengths and weaknesses can be identified more clearly. For example, data integration, visualisation, and predictive analytics could be assessed as separate parts of the framework. The sources informing this requirement are listed in table 4.12.

Table 4.12: Literature evidence for RS3: modularity

Code	Key evidence / citation from source	Relevant paper	Evidence prominence
RS3	The study constructs a qualitative evaluation system comprising 5 dimensions and 13 rubrics, allowing for separate assessment of separable but related components like content and reliability.	(Li et al., 2024)	Very high
RS3	The maturity model is structured such that it can be applied to different levels, starting from a single dimension to the entire digital twin, allowing for identifying specific strengths and weaknesses in each.	(Uhlenkamp et al., 2022)	High

4.3.4. RS4: Operationalisability and scoring clarity

Another structural requirement is operationalisability and scoring clarity (RS4). To be useful in practice, the framework should translate requirements into clearly defined indicators or rubrics that can be assessed consistently across cases and assessors. In other words, the framework should not remain at the level of abstract design principles, but should provide a clear basis for assessment and comparison. The supporting evidence for this requirement is provided in table 4.13.

Table 4.13: Literature evidence for RS4: operationalisability and scoring clarity

Code	Key evidence / citation from source	Relevant paper	Evidence prominence
RS4	The framework uses a D-ANP method to evaluate the maturity level and assign weights to rubrics, transforming qualitative judgment into a quantitative system.	(Li et al., 2024)	Very high
RS4	The maturity model assigns scores to dimensions and attributes, calculates maturity based on achieved versus maximum scores, and applies weighting to reflect the relative importance of categories, thereby supporting a transparent and structured scoring procedure.	(Uhlenkamp et al., 2022)	High
RS4	This research develops assessment rubrics as a core part of the model to make it computable, including evaluative criteria, quality definitions at particular levels, and a scoring strategy, thereby making maturity evaluation operational and repeatable across cases.	(L. Chen et al., 2021)	High

4.3.5. RS5: Balance between technical and organisational maturity

A further structural requirement is a balance between technical and organisational maturity (RS5). A digital twin may be technically advanced, but still difficult to implement or use effectively if the surrounding organisation is not ready for it. For example, even a well-developed system may fail if stakeholders do not share data, if organisational change is not supported, or if users do not trust the system outputs. Table 4.14 shows how this requirement is grounded in the reviewed literature.

Table 4.14: Literature evidence for RS5: balance between technical and organisational maturity

Code	Key evidence / citation from source	Relevant paper	Evidence prominence
RS5	The model is structured around Purpose, Function, and Trust, showing that digital twin maturity depends not only on technical capability but also on organizational, governance, and trust-related conditions.	(L. Chen et al., 2021)	Very high
RS5	The model incorporates two additional dimensions—System and Organizational perspectives—to explicitly assess the effects of non-technical factors alongside the technical ones.	(Z.-S. Chen et al., 2024)	Very high
RS5	High maturity should not imply the exclusion of humans; the model identifies that organizational culture and change management are critical for the successful adoption of new technologies.	(Mo et al., 2023)	High

4.3.6. RS6: Actionable and prioritised progression pathway

Finally, the framework should provide an actionable and prioritised progression pathway (RS6). This means that the assessment should not only indicate the current maturity level, but should also help identify gaps and possible next steps for improvement. The framework should therefore support development guidance rather than only maturity description. The literature basis for this requirement is summarised in table 4.15.

Table 4.15: Literature evidence for RS6: actionable and prioritised progression pathway

Code	Key evidence / citation from source	Relevant paper	Evidence prominence
RS6	The study proposes a stepwise implementation roadmap consisting of development, integration, and advancement phases to guide airports in their digital twin journey.	(Moon et al., 2025)	Very high
RS6	The maturity model provides an actionable roadmap by identifying key gaps and weaknesses across dimensions and subdimensions, and by specifying next steps and a construction trajectory for improving digital twin-supported asset management performance.	(L. Chen et al., 2021)	High

4.4. Environmental Requirements: The Context

Environmental requirements describe the broader conditions under which the maturity assessment framework must be applicable. These requirements ensure that the framework does not focus only on technical capability, but also on the practical conditions that influence whether a digital twin can create value in airport turnaround operations. This is important because turnaround operations take place in a multi-actor environment, where usability, governance, interoperability, security, and trust all affect whether digital twin capabilities can be used in practice.

Table 4.16 provides a compact overview of the literature supporting each environmental requirement. The detailed evidence for each requirement is presented in the separate tables below.

4.4.1. RE1: Usability

One environmental requirement is usability (RE1). The framework should consider whether the digital twin can be used effectively by its intended operational and strategic users. In practice, this means that relevant stakeholders should be able to interact with digital twin outputs in a way that supports their tasks and decisions. The evidence base for this requirement is shown in table 4.17.

Table 4.16: Traceability matrix linking environmental requirements to supporting literature

Requirement	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	P15	P16	P17	P18	P19	P20	P21	P22	P23	P24
RE1		X							X		X													
RE2					X				X						X									
RE3	X	X	X																	X				
RE4	X	X										X												
RE5			X												X			X						
RE6						X			X										X					

Requirement codes: RE1 = Usability; RE2 = Understandability; RE3 = Interoperability; RE4 = Governance compatibility; RE5 = Data security and confidentiality; RE6 = Data and model trustworthiness.

Paper legend: P1 = (L. Chen et al., 2021); P2 = (Moon et al., 2025); P3 = (Conde et al., 2022); P4 = (Attar et al., 2024); P5 = (Van Der Horn & Mahadevan, 2021); P6 = (C. Liu et al., 2024); P7 = (Uhlenkamp et al., 2022); P8 = (Luo et al., 2024); P9 = (Manickam et al., 2023); P10 = (Qiu, 2024); P11 = (K. Chen et al., 2024); P12 = (Haraguchi et al., 2024); P13 = (Z.-S. Chen et al., 2024); P14 = (Li et al., 2024); P15 = (Gourisetti et al., 2025); P16 = (Lu et al., 2025); P17 = (Mo et al., 2023); P18 = (Scheibmeir & Malaiya, 2022); P19 = (Mizrak et al., 2024); P20 = (Hartmann & Wimmer, 2026); P21 = (Saifutdinov et al., 2022); P22 = (Y. Liu et al., 2024); P23 = (Digital Twin Consortium, 2024); P24 = (Madni et al., 2019).

Table 4.17: Literature evidence for RE1: usability

Code	Key evidence / citation from source	Relevant paper	Evidence prominence
RE1	"Usability" is defined as a core pillar of user experience (UX), measuring the ease of use and learnability of the twin in creating meaningful interactions for the operator.	(Manickam et al., 2023)	Very high
RE1	A user study with licensed air traffic controllers (ATCOs) validated that a mixed reality (MR)-based digital twin's visual and interactive design was acceptable and useful for ground control tasks.	(K. Chen et al., 2024)	High
RE1	The study highlights that a lack of user-oriented platforms is a core challenge and identifies that system interfaces must reflect the actual functions needed by maintenance workers.	(Moon et al., 2025)	Moderate

4.4.2. RE2: Understandability

Closely related to usability is understandability (RE2). The framework should consider whether digital twin outputs, indicators, and recommendations are intelligible and interpretable for relevant stakeholders. Even technically advanced outputs have limited practical value if users cannot understand what they mean or how they should inform action. Table 4.18 presents the literature supporting this requirement.

Table 4.18: Literature evidence for RE2: understandability

Code	Key evidence / citation from source	Relevant paper	Evidence prominence
RE2	The "Understandability" pillar focuses on sensemaking, ensuring the user is aware of the scenario, context, and the specific tasks required to fulfill operational objectives.	(Manickam et al., 2023)	High
RE2	"Explainability" is a critical pillar for AI trust; digital twins must enable clear and understandable explanations of how decisions and predictions are derived to be useful in high-stakes environments.	(Gourisetti et al., 2025)	High
RE2	Visualization tools serve to improve the overall acceptance of the digital twin and demonstrate its value to non-domain experts and organizational leadership.	(Van Der Horn & Mahadevan, 2021)	Moderate

4.4.3. RE3: Interoperability

A further environmental requirement is interoperability (RE3). In this study, interoperability refers to the ability of different systems and data sources to work together and use shared information consistently. This is particularly important in airport turnaround operations, where information is spread across multiple actors and systems. Without interoperability, data remain fragmented, limiting the extent to which a digital twin can provide an integrated view of operations. The sources informing this requirement are listed in table 4.19.

Table 4.19: Literature evidence for RE3: interoperability

Code	Key evidence / citation from source	Relevant paper	Evidence prominence
RE3	Mandates the use of OpenBIM and Common Data Environments (CDE) to solve the problem of missing standards and ensure interoperability across various airport software applications.	(Moon et al., 2025)	Very high
RE3	The study utilizes FIWARE and NGSI-LD to manage and homogenize heterogeneous data collected from diverse airport actors and external APIs like Azinq Chroma.	(Conde et al., 2022)	High
RE3	Defines the highest stage of maturity as interoperability across digital twins, where multiple independent twins exchange information to optimize a broader "System of Systems".	(Hartmann & Wimmer, 2026)	Moderate
RE3	The maturity model emphasizes integration of heterogeneous digital products into a semantically rich integrated model, strengthening interoperability beyond technical data exchange alone.	(L. Chen et al., 2021)	Moderate

4.4.4. RE4: Governance compatibility

The framework should also consider governance compatibility (RE4). Airport turnaround operations are shaped by safety rules, accountability structures, and operational governance. The maturity framework should therefore be applicable within these constraints and should support assessment without ignoring existing responsibilities and decision rights. The supporting evidence for this requirement is provided in table 4.20.

Table 4.20: Literature evidence for RE4: governance compatibility

Code	Key evidence / citation from source	Relevant paper	Evidence prominence
RE4	Governance involves accountability and transparency in the decision-making process, especially as twins move toward autonomous reasoning where responsibility must be clearly assigned.	(Haraguchi et al., 2024)	Very high
RE4	Digital twin implementation is strictly limited by compliance with safety and security regulations, requiring the information strategy to be aligned with institutional and regulatory constraints.	(Moon et al., 2025)	High
RE4	The "Curation" principle of the Gemini Principles requires that the digital twin must be clearly owned, governed, and regulated to function effectively within its ecosystem.	(L. Chen et al., 2021)	High

4.4.5. RE5: Data security and confidentiality

A further environmental requirement is data security and confidentiality (RE5). Because digital twins rely on continuous data exchange across actors and systems, the framework should consider whether operational data are handled securely and in a way that protects sensitive or confidential information. This is particularly important in airport environments, where data sharing may involve commercially sensitive, operational, or safety-related information. Table 4.21 shows how this requirement is grounded in the reviewed literature.

Table 4.21: Literature evidence for RE5: data security and confidentiality

Code	Key evidence / citation from source	Relevant paper	Evidence prominence
RE5	Organizations must ensure the quality and security of data through "security by design", including cryptographic integrity checks on data streams and data provenance.	(Gourisetti et al., 2025)	Very high
RE5	Digital twins increase the cyber-attack surface; a maturity model must assess the twin's ability to secure itself and the physical asset through a prioritized cybersecurity roadmap.	(Scheibmeir & Malaiya, 2022)	High
RE5	Managing access for different actors and guaranteeing the security of the system are identified as significant barriers to the successful operation of airport digital twins.	(Conde et al., 2022)	Moderate

4.4.6. RE6: Data and model trustworthiness

Finally, the framework should consider data and model trustworthiness (RE6). This means assessing whether the data, analytics, and outputs of the digital twin are reliable, consistent, and transparent enough to support operational decision-making. In real-time airport settings, this is especially important because decisions may depend on continuously updated information from multiple sources. The literature basis for this final environmental requirement is summarised in table 4.22.

Table 4.22: Literature evidence for RE6: data and model trustworthiness

Code	Key evidence / citation from source	Relevant paper	Evidence prominence
RE6	The "Dependability" pillar measures the twin's ability to be reliable, accurate, and consistent, ensuring users have the confidence to use outputs for decision-making.	(Manickam et al., 2023)	High
RE6	"Missing data and anomalies" are identified as primary limitations in real-time applications, underlining the requirement for data quality and validation mechanisms to support decisions.	(C. Liu et al., 2024)	High
RE6	Incorporating blockchain technology ensures a "transparent and tamper-proof recording" of metrics, which is vital for maintaining compliance and stakeholder trust.	(Mizrak et al., 2024)	Moderate

4.5. Overview of the Requirements

The previous sections presented the functional, structural, and environmental requirements and linked them to the supporting literature. Only sources with moderate, high, or very high evidence prominence were included in the detailed evidence tables. Statements with low evidence prominence were not included, because these only mentioned a concept briefly or indirectly and therefore did not provide sufficient support for deriving a requirement.

Table 4.23 provides a consolidated overview of the final requirement set. The table summarises the requirement codes, requirement names, and short descriptions, and therefore serves as a compact overview of the requirements that will be used in the development of the maturity assessment framework.

Table 4.23: Requirements for the airport turnaround digital twin maturity assessment framework

Code	Requirement	Short description
Functional requirements (what the framework should assess in the digital twin)		
RF1	Real-time operational visibility	The framework should assess whether the digital twin provides a current and continuously updated view of turnaround status, resources, and disruptions.
RF2	Integrated and semantically consistent data foundation	The framework should assess whether the digital twin combines data from multiple airport systems and stakeholders into a coherent, timely, and consistently defined operational representation.
RF3	Proactive operational intelligence	The framework should assess whether the digital twin can identify emerging turnaround problems, diagnose likely causes, and estimate their operational consequences.
RF4	Decision support for turnaround interventions	The framework should assess whether the digital twin can support the selection and evaluation of appropriate operational responses when disruptions occur.
RF5	Resource optimisation	The framework should assess whether the digital twin can support efficient allocation and utilisation of turnaround resources such as staff, vehicles, and stand-related equipment.
RF6	Dynamic rescheduling capability	The framework should assess whether the digital twin can support timely reallocation and rescheduling of activities and resources when operational conditions change.
RF7	Shared situational awareness	The framework should assess whether the digital twin enables relevant stakeholders to work from a common and up-to-date operational picture.
Structural requirements (how the framework should be organised)		
RS1	Multi-level maturity structure	The framework should contain distinct maturity levels that represent meaningful progression in digital twin capability.
RS2	Multi-dimensional assessment structure	The framework should assess maturity across multiple dimensions rather than reduce it to a single overall score only.
RS3	Modularity	The framework should allow separate assessment of different digital twin capabilities or components in order to identify specific strengths and weaknesses.
RS4	Operationalisability and scoring clarity	The framework should translate the functional and environmental requirements into clearly defined indicators or rubrics that can be assessed consistently across cases and assessors.
RS5	Balance between technical and organisational maturity	The framework should assess both technological capabilities and the organisational conditions required for effective use and implementation.
RS6	Actionable and prioritised progression pathway	The framework should not only assess the current maturity level, but also identify key gaps, priorities, and next development steps.
Environmental requirements (conditions for applicability of the framework)		
RE1	Usability	The framework should consider whether the digital twin can be used effectively by its intended operational and strategic users.

Continued on next page

Code	Requirement	Short description
RE2	Understandability	The framework should consider whether digital twin outputs, indicators, and recommendations are intelligible and interpretable for relevant stakeholders.
RE3	Interoperability	The framework should assess whether the digital twin can exchange and use information across systems, actors, and organisational boundaries.
RE4	Governance compatibility	The framework should be applicable within airport constraints related to safety, accountability, regulation, and operational governance.
RE5	Data security and confidentiality	The framework should consider whether operational data are handled securely and in a way that protects sensitive or confidential information.
RE6	Data and model trustworthiness	The framework should consider whether the data, analytics, and outputs of the digital twin are reliable, transparent, and trustworthy enough to support decision-making.

Part C

Design and Develop Artefact

Introduction to Part C

Before moving into the detailed development of the maturity assessment framework, this part first introduces the abstract structure of the final artefact. This high-level version is included to guide the reader through the following chapters and to show how the different components of the framework fit together. The framework is designed as a maturity assessment instrument for digital twins in airport turnaround operations. It does not only assess whether a digital twin exists, but also how mature its capabilities are and which conditions limit or enable further development.

As shown in Figure 4.1, the maturity assessment framework consists of five main components. First, the maturity levels describe the development path from a baseline situation towards a digital model, digital shadow, monitoring twin, predictive twin, and optimising twin. Second, the external implementation constraints describe contextual conditions that may limit the use of digital twin capabilities in practice. Third, the internal maturity dimensions divide digital twin maturity into capability areas, such as data, system integration, prediction, decision support, operational visibility, and human interaction. Fourth, indicators make these dimensions assessable in practice. Fifth, the scoring and threshold logic translates the assessment into an internal maturity level, an external constraint level, and a gap between the two. This gap is used to identify capability gaps and development priorities.

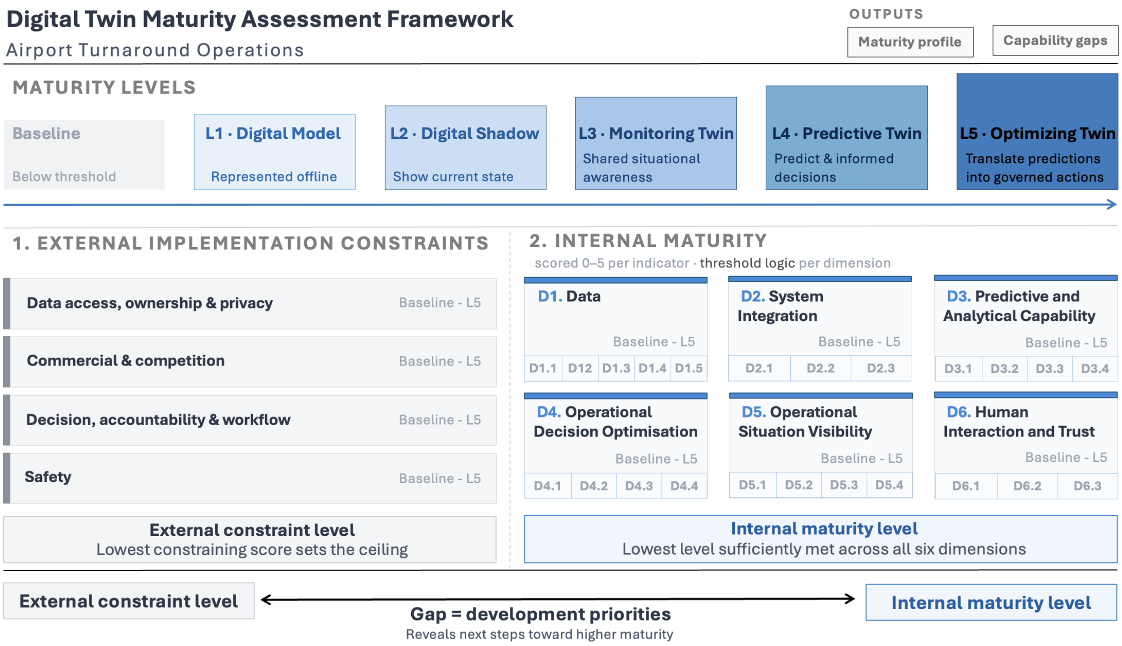


Figure 4.1: Abstract structure of the digital twin maturity assessment framework for airport turnaround operations

Figure 4.1 should be read as a high-level map of the framework. The figure does not yet present the fully detailed framework, but shows the main components that are developed across this part of the thesis. Chapter 5 develops the first literature-based version by translating the requirements into maturity levels, dimensions, indicators, and preliminary scoring logic. Chapter 6 then uses expert feedback to refine these components. Chapter 7 presents the final version of the framework and explains how it can be applied in an airport turnaround context. In this way, Chapters 5–7 gradually fill in the abstract structure introduced here.

5
Literature-Based Framework Development

This chapter presents the initial literature-based maturity assessment framework and forms the first part of the three-step answer to the following sub-question:

SQ2 - Which maturity dimensions, levels, and assessment logic should a digital twin maturity assessment framework for airport turnaround operations contain?

The three-step development process used to answer SQ2 is shown in Figure 5.1.

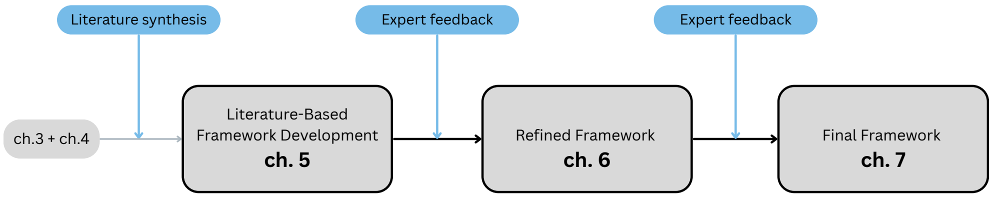


Figure 5.1: Three-step development process of the maturity assessment framework for SQ2

Figure 5.1 shows that SQ2 is answered across three consecutive chapters. In this chapter, the literature synthesis from Chapter 3 and the requirements from Chapter 4 are translated into an initial literature-based framework. In Chapter 6, this initial framework is reviewed and refined through expert feedback. In Chapter 7, the refined framework is reviewed once more and finalised into the maturity assessment framework used in the demonstration and evaluation phases.

The methodological procedure used to develop the initial framework is explained in Chapter 2. Therefore, this chapter focuses on the framework itself. It presents the requirement-to-framework traceability, the literature-based framework with the maturity levels, the dimensions, the indicators, and the preliminary scoring logic.

5.1. Traceability from requirements to framework structure

The requirements defined in Chapter 4 are reflected in the initial framework structure. Table 5.1 shows how each requirement informed one or more framework dimensions or broader design elements. The table is included to make the connection between the requirement set and the resulting framework explicit.

The requirements should not be read as direct one-to-one assessment criteria. Some requirements are mainly reflected in one dimension, while others shape multiple dimensions or framework-wide design choices. For example, requirements related to decision support, resource optimisation, and dynamic rescheduling are mainly reflected in the dimension *Decision support and optimisation*. Requirements related to usability, understandability, and shared situational awareness are reflected more strongly in the dimension *Human interaction and collaboration*.

Table 5.1: Traceability from requirements to framework dimensions and design elements

Code	Requirement	Framework translation	Why this connection is made
RF1	Real-time operational visibility	D5 Operational representation	Relates to whether the digital twin provides a current and continuously updated view of turnaround operations.
RF2	Integrated and semantically consistent data foundation	D1 Data infrastructure; D2 System integration	Requires both a solid data foundation and integration across systems and stakeholders.
RF3	Proactive operational intelligence	D3 Predictive and analytical capability	Relates to the ability to detect, diagnose, and anticipate operational problems.
RF4	Decision support for turnaround interventions	D4 Decision support and optimization	Relates to whether the digital twin supports the evaluation and selection of operational interventions.
RF5	Resource optimisation	D4 Decision support and optimization	Relates to the efficient allocation and utilisation of turnaround resources.
RF6	Dynamic rescheduling capability	D4 Decision support and optimization	Relates to the ability to adapt plans and reallocate resources when conditions change.
RF7	Shared situational awareness	D5 Operational representation; D6 Human interaction and collaboration	Requires both a shared operational picture and its effective use by stakeholders.
RS1	Multi-level maturity structure	Framework-wide: maturity level architecture	Shapes the level structure of the framework as a whole.
RS2	Multi-dimensional assessment structure	Framework-wide: D1–D7	Is reflected in the choice to organise the framework across multiple dimensions.
RS3	Modularity	Framework-wide: D1–D7	Is reflected in the possibility to assess dimensions separately within one coherent framework.
RS4	Operationalisability and scoring clarity	Framework-wide: indicators, rubrics, and scoring logic	Shapes the formulation of indicators and the consistency of scoring across the framework.
RS5	Balance between technical and organisational maturity	Framework-wide, next to D6 Human interaction and collaboration and D7 Governance, interoperability, and trust	Is reflected in the inclusion of both technical and organisational dimensions.
RS6	Actionable and prioritised progression pathway	Framework-wide: interpretation and improvement logic	Shapes how assessment outcomes are translated into priorities and next steps.
RE1	Usability	D6 Human interaction and collaboration	Relates to whether intended users can work with the digital twin effectively.
RE2	Understandability	D6 Human interaction and collaboration	Relates to whether outputs and recommendations are easy for stakeholders to interpret.

Continued on next page

Code	Requirement	Framework translation	Why this connection is made
RE3	Interoperability	D2 System integration; D7 Governance, interoperability, and trust	Covers both technical integration and exchange across organisational boundaries.
RE4	Governance compatibility	D7 Governance, interoperability, and trust	Relates to alignment with accountability, regulation, and operational governance structures.
RE5	Data security and confidentiality	D7 Governance, interoperability, and trust	Relates to the secure handling of sensitive operational data and controlled access to it.
RE6	Data and model trustworthiness	D3 Predictive and analytical capability; D7 Governance, interoperability, and trust	Covers both the reliability of outputs and the confidence needed to use them in decision-making.

5.2. Literature basis of the framework

The framework is supported by three streams of literature. First, general digital twin maturity literature supports the overall maturity progression from basic digital representation towards monitoring, prediction, prescription, optimisation, and, in some cases, autonomy (Kritzinger et al., 2018; Y. Liu et al., 2024; Madni et al., 2019). Second, detailed digital twin maturity models support the use of multiple dimensions, indicators, rubrics, and scoring logic to assess maturity in a structured way (L. Chen et al., 2021; Li et al., 2024; Uhlenkamp et al., 2022). Third, airport- and turnaround-related digital twin literature supports the contextualisation of the framework for airport turnaround operations, including operational visibility, system integration, prediction, rescheduling, resource use, and decision support (Attar et al., 2024; Conde et al., 2022; C. Liu et al., 2024; Luo et al., 2024).

The following sections present how these literature streams support the selected maturity levels, dimensions, indicators, and preliminary scoring logic.

5.3. Maturity levels from literature

The framework distinguishes five maturity levels. These levels describe the overall progression of an airport turnaround digital twin, from a basic digital representation of the turnaround process to a twin that can support optimisation and operational decision-making. The maturity levels provide the main structure of the framework. The dimensions and indicators introduced later make this structure more detailed by specifying which capabilities should be assessed at each level.

The maturity levels are supported by existing digital twin maturity models and related maturity structures. The reviewed literature commonly describes maturity as a progression from basic digital representation towards more advanced capabilities, including monitoring, prediction, prescription, optimisation, integration, and in some cases autonomy. The literature used to support the maturity-level structure is summarised in table 5.2.

Table 5.2: Digital twin maturity levels identified in the literature

Key evidence / description from source	Relevant paper
The study distinguishes three levels of digital twin development based on the degree of automated data exchange: Digital Model, Digital Shadow, and Digital Twin.	(Kritzinger et al., 2018)
The study proposes a five-level maturity progression from Digital Model (L0) to Federated Twin (L4), with increasing capability and shared access.	(Y. Liu et al., 2024)

Continued on next page

Key evidence / description from source	Relevant paper
The study proposes a five-level maturity progression for building digital twins, from BIM-based Digital Model (L1) to fully autonomous Ideal BDT (L5).	(Z.-S. Chen et al., 2024)
The study defines four levels of digital twin maturity: Pre-Digital Twin, Digital Twin, Adaptive Digital Twin, and Intelligent Digital Twin.	(Madni et al., 2019)
The study uses three dimensions, nine sub-dimensions, and 27 rubrics to assess digital twin maturity in a structured way.	(L. Chen et al., 2021)
The study proposes an eight-stage maturity model, ranging from Planning (S1) to Proactive Real-time Synchronisation and Autonomous Implementation (S8).	(Haraguchi et al., 2024)
The study develops a multi-dimensional maturity model with ranked characteristics, using staged levels to assess digital twin development.	(Uhlenkamp et al., 2022)
The report presents airport-specific maturity levels that progress from Descriptive to Diagnostic, Predictive, Prescriptive, and Autonomous capabilities.	(Digital Twin Consortium, 2024)
The study proposes a five-level maturity hierarchy: Description, Reflection, Prediction, Prescription, and Fusion.	(Li et al., 2024)
The literature comparison shows that digital twin maturity models commonly describe a progression from basic representation toward more integrated, intelligent, and autonomous twin capabilities.	(Hartmann & Wimmer, 2026)

As shown in table 5.2, several reviewed maturity models use autonomy as the highest level. In these models, the digital twin does not only monitor, analyse, or predict the physical system, but can also influence or control it automatically. This can be useful in contexts such as automated machines, building systems, urban infrastructure, or system-wide control, where automatic action is part of the expected use of the digital twin (Z.-S. Chen et al., 2024; Digital Twin Consortium, 2024; Haraguchi et al., 2024; Y. Liu et al., 2024). For airport turnaround operations, however, full autonomous execution is less suitable because the process is safety-critical, time-sensitive, and divided across multiple responsible actors, such as airlines, ground handlers, airport operators, and air traffic-related stakeholders. In this context, the digital twin should support human decision-making by providing timely information, predictions, recommendations, and intervention options, rather than independently executing operational actions. This is consistent with literature that emphasises human-in-the-loop decision support, explainability, trust, and safe use of digital twin outputs in complex operational environments (K. Chen et al., 2024; Uhlenkamp et al., 2022; Van Der Horn & Mahadevan, 2021).

Therefore, the highest maturity level in this framework focuses on advanced decision support rather than autonomous control. At this level, the digital twin can recommend actions and support operational interventions, but final decisions and actions remain with the responsible human actors and organisations. This makes the maturity structure more suitable for the airport turnaround context, where coordination, accountability, and safety are more important than full autonomous execution. The resulting five maturity levels used in this research are presented in table 5.3.

Table 5.3: Literature-based maturity levels for airport turnaround digital twins

Level	Name	Description
Level 1	Digital model	The turnaround process is represented digitally, but the model is largely static or manually updated. It mainly supports documentation and basic visualisation.
Level 2	Digital shadow	Operational data flow automatically from the physical turnaround environment to the digital environment. The system reflects operations in near real time, but interaction remains largely one-way.

Continued on next page

Level	Name	Description
Level 3	Monitoring twin	The digital twin provides integrated monitoring of aircraft status, task progress, delays, and resource usage. The system supports operational awareness across multiple turnaround actors.
Level 4	Predictive twin	The twin supports diagnosis, forecasting, and scenario analysis. It helps identify emerging problems and estimate likely operational consequences.
Level 5	Optimising twin	The twin supports prescriptive decision-making through dynamic rescheduling, resource optimisation, and recommendation of operational responses to disturbances.

5.4. Literature basis for the dimensions and indicators

The maturity levels provide the overall progression of the framework, but they do not show in detail which aspects of the digital twin should be assessed. Therefore, seven framework dimensions were defined. These dimensions are supported by both general digital twin maturity literature and airport-specific or turnaround-specific digital twin studies.

The general maturity-model literature was used to structure the dimensions. Uhlenkamp et al. (2022) support the use of separate categories for data, integration, models, control, and human-machine interaction. Li et al. (2024) support dimensions related to data richness, model completeness, prediction, and prescription. L. Chen et al. (2021) provide support for governance, trust, ownership, and the broader purpose of the digital twin.

Airport and turnaround-specific literature was used to make the dimensions suitable for the airport turnaround context. Conde et al. (2022) support the need for system integration and operational representation across different turnaround actors. C. Liu et al. (2024) support the predictive and analytical dimension through delay diagnosis and delay propagation analysis. Luo et al. (2024) and Attar et al. (2024) support the decision support and optimisation dimension through rescheduling and scenario-based analysis. K. Chen et al. (2024), Manickam et al. (2023), and Digital Twin Consortium (2024) support the human interaction and collaboration dimension. Finally, Moon et al. (2025), Gourisetti et al. (2025), and Haraguchi et al. (2024) support the governance, interoperability, and trust dimension.

Together, these sources show that maturity should not be assessed as one single score. Instead, the framework should assess several dimensions, because a digital twin can be mature in one area, such as data infrastructure, while still being less mature in another area, such as governance or decision support. The literature evidence used to define each dimension is presented in the tables below.

Each dimension is further specified through a set of indicators. These indicators were developed by interpreting the reviewed literature and translating the broader requirements into more concrete assessment points. Because the airport turnaround digital twin context is still an emerging research area, the indicators should be understood as an initial literature-based proposal. In the next phase of the research, they will be discussed with experts who have specific knowledge of airport turnaround operations. This expert input will be used to assess whether the indicators are understandable, relevant, and suitable for practical use.

5.4.1. D1: Data Infrastructure

This dimension assesses the quality of the data foundation on which the turnaround digital twin is built. It mainly focuses on the underlying digital twin platform. It concerns whether relevant operational data are available, captured in a sufficiently automatic way, updated frequently enough, and checked for completeness and quality. It also considers whether the digital twin has access to both current and historical operational data, so that turnaround processes can be represented consistently over time. In other words, this dimension is not about what the digital twin can infer, predict, or optimise, but about whether the underlying data basis is timely, reliable, and complete enough to support these functions.

The indicators belonging to this dimension are:

- **D1.1 Automatic data capture:** Assesses whether turnaround data are collected automatically from relevant operational sources rather than entered manually.
- **D1.2 Timely data updating:** Assesses whether operational data are updated frequently enough to reflect the current turnaround situation.
- **D1.3 Data completeness:** Assesses whether the digital twin includes sufficiently complete data on relevant turnaround activities, resources, and milestones.
- **D1.4 Data quality management:** Assesses whether data quality issues, such as inconsistency, inaccuracy, or missingness, are identified and handled systematically.
- **D1.5 Integration of live and historical data:** Assesses whether real-time and historical data are available together to support both current awareness and longer-term analysis.

The literature evidence supporting this dimension is presented in table 5.4.

Table 5.4: Literature evidence for D1: Data infrastructure

Key evidence / description from source	Relevant paper
Digital twin implementation requires communication infrastructure that can support continuous and high-frequency data exchange between the physical and virtual system. Technologies such as wireless sensors, RFID, IoT, and cloud computing can support flexible data collection, storage, and processing.	(Van Der Horn & Mahadevan, 2021)
Airport digital twins can use IoT-based sensing and low-power wide-area network technologies, such as LoRaWAN, to collect and transmit data across large airport facilities. These technologies support real-time data capture from distributed devices and can be connected to existing operational control systems.	(Durmaz & Göçmen, 2025)
Data is included as a separate maturity category in the maturity model. This category assesses aspects such as data storage, data scope, data sources, and data quality, which supports the inclusion of data infrastructure as a separate framework dimension.	(Uhlenkamp et al., 2022)

5.4.2. D2: System Integration

This dimension assesses whether the digital twin can connect and combine the different systems, data sources, and organisational boundaries that structure airport turnaround operations. It mainly focuses on the underlying digital twin platform. In a turnaround context, relevant information is typically distributed across airport systems, airline systems, handler systems, and service-provider systems. A more mature twin therefore requires more than isolated data feeds. It requires shared identifiers, compatible interfaces, and a coherent structure through which information can be exchanged and interpreted across actors. This dimension is therefore concerned with the degree to which the twin functions as an integrated operational environment rather than as a collection of disconnected tools.

The indicators belonging to this dimension are:

- **D2.1 Multi-stakeholder data integration:** Assesses whether data from multiple turnaround stakeholders are integrated into the digital twin environment.
- **D2.2 Shared operational view:** Assesses whether different operational systems are combined into one shared and coherent operational view.
- **D2.3 Shared identifiers and interfaces:** Assesses whether common identifiers, data structures, or interfaces are used to connect different systems and sources.
- **D2.4 Cross-boundary information exchange:** Assesses whether information can be exchanged across organisational boundaries with limited manual re-entry or duplication.

The literature evidence supporting this dimension is presented in table 5.5.

Table 5.5: Literature evidence for D2: System integration

Key evidence / description from source	Relevant paper
Airport digital twin architectures need to manage and align heterogeneous data from different airport actors, such as airlines, ground handlers, and airport operators. Open standards such as FIWARE and NGSI-LD can support context management and interoperable airport data models, allowing information to be exchanged and interpreted across organisational boundaries.	(Conde et al., 2022)
A mature airport digital twin can be developed through a system-of-systems perspective, where the airport is viewed as a set of interconnected systems rather than separate silos. A Common Information Model and a Common Data Environment can help connect different operational domains, such as baggage handling, flight planning, and ground operations.	(Digital Twin Consortium, 2024)
System integration is treated as a core maturity category. It concerns the automated connection between different entities and their underlying source systems, and it can develop from a single-unit level towards a broader system-of-systems level.	(Uhlenkamp et al., 2022)

5.4.3. D3: Predictive and Analytical Capability

This dimension assesses whether the digital twin can use available data to generate insight beyond simple status display. It mainly focuses on the underlying digital twin platform. It includes the ability to detect emerging problems, estimate the likely effects of disruptions, explore alternative scenarios, and use current and historical data to produce diagnoses or forecasts. The main question in this dimension is therefore whether the twin supports explanation and anticipation, rather than only observation. This is especially relevant in airport turnaround operations, where delays and disturbances develop dynamically and where value increasingly lies in identifying problems before they fully materialise.

The indicators belonging to this dimension are:

- **D3.1 Early problem detection:** Assesses whether the digital twin can identify emerging turnaround problems before they fully materialise.
- **D3.2 Impact estimation:** Assesses whether the digital twin can estimate the likely operational effects of disruptions or deviations.
- **D3.3 Scenario analysis:** Assesses whether the digital twin supports what-if analysis or exploration of alternative operational scenarios.
- **D3.4 Data-driven prediction:** Assesses whether the digital twin uses current and historical data to generate predictions about turnaround performance.
- **D3.5 Causal analysis support:** Assesses whether the analytical outputs help users distinguish likely causes of problems from visible symptoms.

The literature evidence supporting this dimension is presented in table 5.6.

Table 5.6: Literature evidence for D3: Predictive and analytical capability

Key evidence / description from source	Relevant paper
A mature airport digital twin can use Bayesian networks to model dynamic relationships between ground service processes and identify likely causes of delay. The study also uses breakpoint simulation to predict delay propagation from an intermediate operational moment, rather than only from the start of the turnaround process.	(C. Liu et al., 2024)

Continued on next page

Key evidence / description from source	Relevant paper
The maturity hierarchy describes a progression from a reflection twin, which shows what is happening, towards a prediction twin, which estimates what is likely to happen next. This supports the inclusion of predictive capability as a separate dimension of digital twin maturity.	(Li et al., 2024)
Higher maturity levels include cognitive digital twin capabilities, where the system can support analysis, simulation, forecasting, and scenario validation. This shows that advanced maturity goes beyond monitoring and includes analytical and predictive functions.	(Y. Liu et al., 2024)

5.4.4. D4: Decision Support and Optimisation

This dimension assesses whether the digital twin translates insight into support for operational action. It mainly focuses on the underlying digital twin platform, but specifically on the way this platform supports operational decision-making. Whereas the previous dimension focuses on understanding and predicting the situation, this dimension focuses on helping users decide what to do about it. It therefore concerns whether the twin can compare response options, support the adjustment of tasks and resources, assist with replanning, and provide recommendations that are operationally usable. A mature digital twin in this sense does not only diagnose or forecast disruption, but also helps actors select and coordinate the most appropriate response under changing conditions.

The indicators belonging to this dimension are:

- **D4.1 Response selection support:** Assesses whether the digital twin supports the selection of operational responses when disruptions or deviations occur.
- **D4.2 Resource adjustment support:** Assesses whether the digital twin supports the reassignment of staff, vehicles, stands, or other resources when conditions change.
- **D4.3 Task and schedule adjustment support:** Assesses whether the digital twin supports changes in the order or timing of turnaround tasks in response to operational developments.
- **D4.4 Comparison of response options:** Assesses whether alternative operational responses and their likely consequences can be compared systematically.
- **D4.5 Actionable recommendations:** Assesses whether the digital twin provides recommendations or priorities that can be directly used in operational practice.

The literature evidence supporting this dimension is presented in table 5.7.

Table 5.7: Literature evidence for D4: Decision support and optimisation

Key evidence / description from source	Relevant paper
A prescription twin goes beyond prediction by supporting the question of what should be done. It uses knowledge-based and behavioural models to provide optimised decision-making support. This supports the inclusion of decision support and optimisation as a higher maturity capability.	(Li et al., 2024)
Digital twins can support event-driven rescheduling when disruptions occur, such as flight delays or vehicle malfunctions. By linking optimisation algorithms with simulation models, the twin can generate updated operational plans and support the coordination of airport ground resources under changing conditions.	(Luo et al., 2024)
Scenario-based what-if analysis can be used to evaluate the likely effects of decisions before they are implemented. This allows managers to compare possible response options and estimate their effects on airport performance indicators such as capacity and delay.	(Attar et al., 2024)

5.4.5. D5: Operational Representation

This dimension assesses how well the digital twin represents the actual turnaround operation itself. It concerns both the underlying operational representation and, partly, how this representation becomes visible through the dashboard or interface. The dimension assesses whether the twin captures the current state of the aircraft, the status and sequencing of turnaround tasks, the use of relevant resources, and the presence and effects of disruptions. It also concerns whether the represented process contains enough operational detail to support understanding and decision-making. The term operational representation is used deliberately to emphasise the quality and completeness of what is represented in the twin, rather than only the ease with which users interact with that representation.

The indicators belonging to this dimension are:

- **D5.1 Current turnaround status visibility:** Assesses whether the digital twin clearly shows the current status of aircraft, turnaround tasks, and relevant resources.
- **D5.2 Task dependency visibility:** Assesses whether the digital twin clearly shows how turnaround activities depend on and affect one another.
- **D5.3 Disruption visibility:** Assesses whether delays, disruptions, and abnormal operational events are clearly represented.
- **D5.4 Operational level of detail:** Assesses whether the representation contains enough operational detail to support understanding and decision-making.

The literature evidence supporting this dimension is presented in table 5.8.

Table 5.8: Literature evidence for D5: Operational representation

Key evidence / description from source	Relevant paper
A digital twin should represent a specific physical system and capture its actual states over time. A mature representation can include information about structure, performance, maintenance history, and health status, which helps users understand how the physical system is performing in practice.	(Madni et al., 2019)
Operational representation depends on model completeness. This can develop from basic attribute and geometric models towards behaviour, physics, and interaction models, allowing the twin to describe not only what an object is, but also how it functions and interacts with its environment.	(Li et al., 2024)
In the airport context, the digital twin should represent the sequence of turnaround events from arrival at the stand to departure. This requires data from multiple actors to show the status of individual tasks, such as refuelling, catering, and baggage handling, and their dependencies within the operational area.	(Conde et al., 2022)

5.4.6. D6: Human Interaction and Collaboration

This dimension assesses whether the dashboard or interface enables users to work with the digital twin effectively in operational practice. It mainly focuses on the user-facing part of the digital twin. It includes the understandability of the interface, its usability in time-sensitive work settings, the extent to which users trust the presented information, and the extent to which the twin supports shared situational awareness and coordination across actors. This dimension is included because even a technically advanced twin has limited practical value if users cannot interpret it easily, rely on it, or use it collaboratively. Maturity therefore also depends on whether the twin functions as a usable and meaningful decision environment for its intended users.

The indicators belonging to this dimension are:

- **D6.1 Understandability:** Assesses whether the dashboard or interface is understandable for its intended users.

- **D6.2 Usability in practice:** Assesses whether the interface can be used effectively in day-to-day turnaround work.
- **D6.3 Shared situational awareness:** Assesses whether the interface supports a shared understanding of the turnaround situation across different users and actors.
- **D6.4 Information trustworthiness:** Assesses whether users consider the presented information sufficiently trustworthy for operational interpretation and decision-making.
- **D6.5 Collaboration support:** Assesses whether the interface supports collaboration and coordination between turnaround actors.

The literature evidence supporting this dimension is presented in table 5.9.

Table 5.9: Literature evidence for D6: Human interaction and collaboration

Key evidence / description from source	Relevant paper
Digital twin user experience maturity can be assessed through understandability, usability, convenience, dependability, and delight. This shows that a mature digital twin should not only provide information, but should also help users understand and use that information in a clear and meaningful way.	(Manickam et al., 2023)
In safety-critical airport ground tasks, shared visualisation and tangible interfaces can support shared situational awareness between multiple operators. By providing a common view of operational landmarks and conflict alerts, the digital twin can help stakeholders coordinate tasks and resolve conflicts with less verbal communication.	(K. Chen et al., 2024)
A mature digital twin can act as a shared platform where stakeholders visualise, understand, and act on data-driven insights together. In the airport context, integrated operations centres can help reduce operational silos and improve coordination across the airport ecosystem.	(Digital Twin Consortium, 2024)

5.4.7. D7: Governance, Interoperability, and Trust

This dimension assesses whether the digital twin is embedded in organisational and institutional arrangements that allow it to function sustainably and legitimately in practice. It focuses on the overall digital twin setup rather than only the technical platform or dashboard. It concerns the clarity of roles and responsibilities, the existence of formal rules for data sharing and data use, the treatment of privacy, security, and access rights, and the alignment of the twin with existing operational, safety, and accountability structures. It also includes the question of whether the twin is designed in a way that supports future interoperability and coordinated development. This dimension is broader than technical integration alone: it addresses whether the digital twin is governable, trusted, and maintainable within a multi-actor airport environment.

The indicators belonging to this dimension are:

- **D7.1 Clarity of roles and responsibilities:** Assesses whether roles, responsibilities, and ownership arrangements related to the digital twin are clearly defined.
- **D7.2 Formalised data-sharing rules:** Assesses whether rules for data sharing, data use, and related governance procedures are formally established.
- **D7.3 Privacy, access control, and security:** Assesses whether appropriate measures for access control, privacy protection, and data security are defined and embedded.
- **D7.4 Alignment with operational and accountability structures:** Assesses whether the digital twin setup is aligned with airport operational rules, safety requirements, and accountability arrangements.
- **D7.5 Future interoperability:** Assesses whether the digital twin is structured in a way that enables future integration with other airport systems and stakeholders.

The literature evidence supporting this dimension is presented in table 5.10.

Table 5.10: Literature evidence for D7: Governance, interoperability, and trust

Key evidence / description from source	Relevant paper
Digital twin maturity can be structured around the Gemini Principles: purpose, trust, and function. This shows that maturity does not only depend on technical capability, but also on clear ownership, governance, regulation, security, openness, and data quality.	(L. Chen et al., 2021)
In complex multi-actor environments, trust can be supported through security-by-design and privacy-by-design approaches. This includes data integrity checks, data provenance, and compliance with privacy requirements, which are important when stakeholders share sensitive operational data.	(Gourisetti et al., 2025)
Governance for large-scale digital twins includes accountability, transparency, and inclusion. A mature system requires clear boundaries and responsibilities between actors, especially when digital twin outputs or AI-supported insights are used in decision-making.	(Haraguchi et al., 2024)
Airport infrastructure requires standardised and interoperable systems to manage complex assets and support collaboration between stakeholders such as airport authorities, airlines, and ground handlers. A structured Information Strategy Plan and Common Data Environment can help address fragmented platforms and inconsistent data standards.	(Moon et al., 2025)

5.5. Initial assessment and scoring logic

This section presents the initial scoring logic of the literature-based framework. The scoring logic is included as part of the initial artefact, but it is not treated as final at this stage. In the next phase, expert feedback is used to assess whether the indicators and thresholds are understandable, relevant, and practically usable.

The main purpose of this chapter was to define the maturity levels, dimensions, and indicators of the framework. The assessment and scoring logic is therefore only included as a preliminary draft. This draft may be discussed with experts if time allows, but it is not yet treated as a final part of the framework. Before finalising the scoring logic, it must first be checked whether experts consider the indicators suitable for scoring on a five-point scale.

The draft scoring approach uses an indicator-based threshold logic. This means that each maturity level is linked to a set of key indicators that need to reach a minimum score before that level can be assigned. This approach was chosen because it makes the assessment more transparent than assigning one overall maturity score only. It also helps prevent a high score in one area from compensating for a weak score in an essential area. For example, a digital twin should not be classified as a predictive twin if the indicators related to prediction and impact estimation are still weak.

Similar threshold-based reasoning can be found in existing digital twin maturity literature. Li et al. (2024) use rubric-level requirements and a "bucket effect", where the maturity level of a dimension is constrained by the performance of its underlying rubrics. L. Chen et al. (2021) use rubric scores and total score ranges to assign maturity stages in their Gemini Principles-based maturity model. Z.-S. Chen et al. (2024) also use a numerical maturity assessment approach in which maturity is linked to defined score intervals.

In this draft, each indicator is assumed to be scored on a five-point scale, where a higher score indicates a more mature implementation. The thresholds in table 6.10 indicate which minimum indicator scores would be needed to assign each maturity level. However, these thresholds are only a preliminary proposal. Expert feedback is needed to determine whether the five-point scale and the proposed threshold values are suitable in practice.

The proposed threshold logic is shown in table 6.10. Level 1 has no formal threshold because it represents the existence of a basic digital representation. From Level 2 onwards, the levels require minimum scores on selected indicators that are considered essential for that level of maturity.

Table 5.11: Literature-based scoring logic for assigning maturity levels

Level	Label	Main capability	Key indicators and threshold
1	Digital model	Turnaround is represented digitally	No formal threshold; basic digital representation exists.
2	Digital shadow	Operational data automatically update the system	**D1.1** Automatic data capture ≥ 3; **D1.2** Timely data updating ≥ 3; **D1.3** Data completeness ≥ 2; **D2.1** Multi-stakeholder data integration ≥ 2
3	Monitoring twin	Shared monitoring across multiple actors	**D1.3** Data completeness ≥ 3; **D2.1** Multi-stakeholder data integration ≥ 3; **D2.2** Shared operational view ≥ 3; **D6.3** Shared situational awareness ≥ 3
4	Predictive twin	Forecasting and impact estimation	**D3.1** Early problem detection ≥ 3; **D3.2** Impact estimation ≥ 3; **D3.3** Scenario analysis ≥ 3; **D3.4** Data-driven prediction ≥ 3; **D3.5** Causal analysis support ≥ 3
5	Optimising twin	Decision support and dynamic optimisation	**D4.1** Response selection support ≥ 4; **D4.2** Resource adjustment support ≥ 4; **D4.3** Task and schedule adjustment support ≥ 4; **D4.4** Comparison of response options ≥ 4; **D4.5** Actionable recommendations ≥ 4

5.6. Full literature-based maturity assessment framework

The literature-based maturity assessment framework forms the starting point of the iterative artefact development process. It combines the maturity levels, dimensions, indicators, and preliminary scoring logic developed in this chapter into one complete initial framework. This full literature-based version is presented in Appendix C.

The full framework is placed in the appendix rather than repeated in the main text. This choice was made to reduce repetition, because the same framework structure is used in the next chapter to present and discuss the expert feedback. In that chapter, the literature-based framework is revisited to show where the feedback applies and how it informed the refinement of the framework.

The literature-based framework was also used as the structure for the expert interviews. By discussing the maturity levels, dimensions, indicators, and scoring logic with experts, it was possible to assess whether the initial framework was understandable, relevant, and suitable for the airport turnaround context. The expert feedback forms the basis for the refined maturity assessment framework presented in the next chapter.

6
Refined Framework: Feedback on Literature-Based Framework

This chapter forms the second part of answering the following sub-question:

SQ2 - Which maturity dimensions, levels, and assessment logic should a digital twin maturity assessment framework for airport turnaround operations contain?

The previous chapter developed the initial literature-based maturity assessment framework from the literature and requirements. This chapter uses expert feedback to review and refine that initial framework. As shown in Figure 6.1, this chapter therefore forms the transition from the literature-based framework to the refined framework.

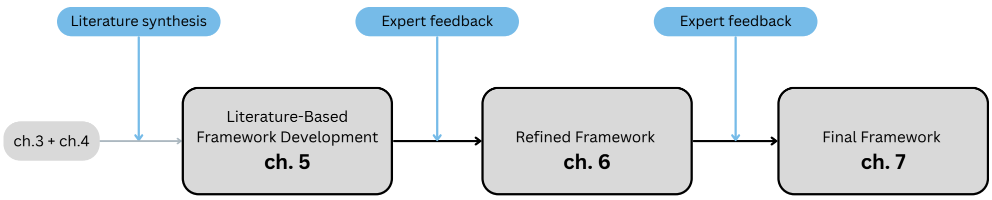


Figure 6.1: Development process of the maturity assessment framework repeated from Figure 5.1

Figure 6.1 is repeated to clarify the position of this chapter within the overall artefact development process. The figure shows that the literature-based framework developed in Chapter 5 is reviewed through expert feedback in this chapter. This feedback is used to refine the maturity levels, dimensions, indicators, external implementation constraints, and scoring logic.

The methodological setup of the interviews and focus group is described in Chapter 2. Therefore, this chapter only briefly introduces the participants and focuses on the feedback and resulting changes to the framework.

6.1. Interview setup

The literature-based framework from Chapter 5 was used as the basis for the expert feedback. The complete version of this framework is provided in Appendix B. To make the refinement process traceable, each feedback section briefly summarises the part of the literature-based framework that was discussed with the experts before explaining the resulting refinement. The literature-based maturity levels and indicators are therefore not repeated in full in this chapter, because the focus is on the feedback and the refined framework. The preliminary threshold logic is repeated, because it was reviewed directly by the experts and forms the basis for the refinement of the scoring logic.

Feedback was collected from five relevant perspectives: academic maturity-model knowledge, digital twin consultancy experience, the airline perspective, the Schiphol operational and digital twin context, and a broader aviation-sector perspective. Together, these perspectives were used to assess whether the maturity levels, dimensions, indicators, and scoring logic were understandable, relevant, and applicable to airport turnaround operations. An overview of the participants and their role in the framework refinement phase is provided in Table 6.1. The expert feedback is reported using interview and participant codes. The interview materials, including the interview overview and semi-structured interview guide, are provided in Appendix C.

Table 6.1: Overview of participants in the framework refinement phase

Interview code	Participant code	Perspective	Main contribution
I1	R	Researcher in digital twins and maturity assessment	Reviewed the conceptual structure of the literature-based framework, including the maturity levels, dimensions, indicators, and scoring logic.
I2	C1	Digital twin consultant	Provided feedback from a digital twin implementation and consultancy perspective.
I3	S1–S3	Schiphol participants	Provided the airport operational and digital twin application perspective, with attention to Wilbur, data use, system integration, governance, trust, and practical assessment.
I4	A1–A2	KLM participants	Provided the airline perspective on turnaround operations, data availability, data sharing, stakeholder dependencies, and airline-related data constraints.
I5	C2	Aviation consultant	Provided an external aviation-sector perspective, with attention to the broader airport context, practical implementation, governance, automation, and the applicability of the assessment in a workshop setting.

6.2. Feedback on the maturity levels

The interviewees reviewed the literature-based maturity levels presented in Appendix B. These levels described a development path from a static digital model, to a near-real-time digital shadow, an integrated monitoring twin, a predictive twin, and finally an optimizing twin. The literature-based table is not repeated here, because this section focuses on how the expert feedback led to the refined maturity levels.

Overall, the maturity levels were considered a logical starting point. R indicated that the first levels were clear and that the level names helped explain the development path:

> *"The first two were totally fine. [...] I think also the names help to understand, so monitoring, predicting, optimizing."* (I1, R)

The main refinement concerned the distinction between Level 2, *Digital Shadow*, and Level 3, *Monitoring Twin*. In the literature-based version, both levels involved digital or real-time information, which made their boundary unclear. C2 summarised this issue as follows:

> *"The levels seem logical. My only question is the difference between Level 2 and Level 3, because in both cases digital data are visible."* (I5, C2, translated)

The refined framework therefore separates a simple current-state view from an integrated monitoring view. Level 2 shows the current operational state in real or near-real time, while Level 3 combines live information from multiple turnaround actors into one shared operational view. The step from Level 2 to Level 3 is therefore not only about real-time data, but about integration and shared situational awareness.

A second refinement concerned the relation between Level 4, *Predictive Twin*, and Level 5, *Optimizing Twin*. In the Schiphol focus group, S2 argued that optimisation requires some ability to look ahead:

> *"To optimize, you need to be able to look ahead. Optimizing only on the basis of what you know right now is not enough."* (I3, S2, translated)

This supported the decision to keep prediction before optimisation. In the refined model, Level 4 focuses on predicting future states and likely operational consequences, while Level 5 focuses on translating these predictions into operational actions, such as task resequencing, resource reassignment, or schedule adjustment.

The interviews also showed that the wording of the levels needed to be more precise. Terms such as "resource usage" could be interpreted in different ways and therefore affect scoring. The refined levels therefore include a clearer role for the digital twin at each stage: *represent offline*, *show current state*, *create shared situational awareness*, *predict and inform decisions*, and *translate predictions into action*.

Finally, the feedback showed that full automated decision-making should not be added as a separate final level. R questioned whether Level 5 implied automatic operational influence, while C2 suggested that automation could be treated as a possible advanced capability within Level 5 rather than as a separate end state (I1, R; I5, C2). The KLM interview further showed that simulation tools and decision engines are not automatically digital twins when they are not connected to real-time operational use (I4, A1–A2).

Overall, the feedback did not lead to removing any maturity level. Instead, it led to sharper definitions, a clearer distinction between *Digital Shadow* and *Monitoring Twin*, a clearer sequence from prediction to optimisation, and the decision to treat automation as a possible capability within Level 5. The refined maturity levels are presented in Table 6.2.

Table 6.2: Refined maturity levels after expert feedback

Level	**Name**	**Clear role of the digital twin**	**Description**
Level 1	Digital Model	Represent offline	A static or non-real-time digital version of the turnaround process. It is mainly used for documentation, offline analysis, or internal understanding. It does not reflect or influence live operations.
Level 2	Digital Shadow	Show current state	A real-time or near-real-time overview of the turnaround operation. The system shows the current operational state, but this view is relatively simple and not yet connected across different turnaround stakeholders. Information mainly flows one way: from the physical operation to the digital system.
Level 3	Monitoring Twin	Create shared situational awareness	A digital twin that combines live information from multiple turnaround actors into one shared view. It shows what is happening now and makes the actions and progress of different actors visible, such as aircraft status, task progress, delays, and resource availability. Operational decisions and actions remain with the existing actors and workflows.
Level 4	Predictive Twin	Predict and inform decisions	A digital twin that helps users understand what is likely to happen next. It uses live and historical data to predict future states, identify risks, and estimate possible consequences for the operation. It informs decision-making, but does not yet translate predictions into structured operational actions.

Continued on next page

Level	Name	Clear role of the digital twin	Description
Level 5	Optimizing Twin	Translate predictions into action	A digital twin that uses predictive insights to identify and prioritise operational decisions, such as task resequencing, resource reassignment, or schedule adjustment. It may propose or initiate decisions through connected systems, while physical execution remains with the responsible actors within defined decision, workflow, accountability, and safety boundaries.

6.3. Feedback on D1 Data infrastructure

The interviewees reviewed the literature-based D1 indicators presented in Appendix B. These indicators assessed whether the digital twin has a sufficient data foundation for turnaround operations, including automatic data capture, timely data updating, data completeness, data quality management, and the combined use of live and historical data. The literature-based table is not repeated here, because this section focuses on how the expert feedback led to the refined indicators.

Overall, D1 Data infrastructure was confirmed as a necessary foundation of the framework. C2 described the sequence from the dimensions data infrastructure to system integration, prediction, and decision support as logical:

> *"D1 to D4 are very logical, because for D2 you also need D1, and for D4 you need D3, D2, and D1."* (I5, C2, translated)

C1 similarly emphasised that the data foundation is the starting point of a digital twin journey, because unreliable or incomplete input data will also lead to unreliable digital twin outputs (I2, C1). Therefore, D1 was retained as the first internal maturity dimension.

The main refinement concerned D1.2, originally called *Timely data updating*. R noted that this wording was unclear, because it did not specify where the updating takes place:

> *"Where is the data being updated? Is the data updated in the digital twin, or in the system around it?"* (I1, R)

The indicator was therefore renamed *Timely data availability*. This better reflects the practical purpose of the indicator: assessing whether captured data become available quickly enough to support the intended operational use.

A second refinement concerned D1.5. In the literature-based version, this indicator combined live and historical data. R argued that live data were already partly covered by automatic data capture and timely data availability. The indicator was therefore sharpened towards historical operational data. This also matched the Schiphol focus group, where S1 noted that:

> *"Historical and live data really depend on the users."* (I3, S1, translated)

In the refined framework, D1.5 therefore assesses whether historical turnaround data are stored and used for learning, prediction, and improvement.

The interviews also showed that data completeness should be interpreted in relation to external implementation constraints. A1 and A2 explained that some airline and ground-operation data may be limited by commercial sensitivity, competition, contractual restrictions, and privacy regulations such as GDPR (I4, A1–A2). This means that a low score on D1.3 or D1.5 may reflect contextual constraints rather than only technical immaturity.

Overall, D1 was retained, but the wording of several indicators was clarified. D1.1, D1.3, and D1.4 were kept with minor wording changes. D1.2 was renamed *Timely data availability*, and D1.5 was renamed *Historical operational data*. The refined D1 indicators are shown in Table 6.3.

Table 6.3: Refined D1 data infrastructure indicators after expert feedback

Code	Indicator	Indicator definition
D1 Data infrastructure		
D1.1	Automatic data capture	Assesses whether relevant turnaround data are captured automatically from operational sources.
D1.2	Timely data availability	Assesses whether captured data become available quickly enough for the intended operational use.
D1.3	Data completeness	Assesses whether the digital twin contains the core turnaround data needed for representation, prediction, and decision support.
D1.4	Data quality management	Assesses whether data quality issues are systematically identified and addressed.
D1.5	Historical operational data	Assesses whether historical turnaround data are stored and used for learning, prediction, and improvement.

6.4. Feedback on D2 System integration

The interviewees reviewed the literature-based D2 indicators presented in Appendix B. These indicators assessed whether the digital twin could integrate data from multiple turnaround stakeholders, combine systems into a shared operational view, use shared identifiers and interfaces, and support information exchange across organisational boundaries. The literature-based table is not repeated here, because this section focuses on how the expert feedback led to the refined indicators.

Overall, D2 System integration was confirmed as a necessary dimension of the framework. R described data infrastructure and system integration as foundational for any digital twin:

> *"Data infrastructure and system integration are classical elements that you always need, regardless of the use case."* (I1, R)

C2 also described D1 to D4 as a logical sequence, in which data infrastructure enables system integration, which then supports prediction and decision support (I5, C2). Therefore, D2 was retained as the second internal maturity dimension.

The main refinement concerned the distinction between D2.1 and D2.4. In the Schiphol focus group, D2.1 was interpreted as the integration of data from actors directly involved in the turnaround, while D2.4 was interpreted as information exchange beyond the direct turnaround scope. S2 explicitly questioned this boundary:

> *"What is then exactly the difference between indicator 1 and indicator 4?"* (I3, S2, translated)

This showed that D2.4 was broader and harder to assess within the internal maturity model. Therefore, D2.1 was sharpened towards direct turnaround stakeholder data, while D2.4 was removed from the internal indicators and treated as part of the external implementation context.

A second refinement concerned the wording of D2.2 and D2.3. R noted that the meaning of a shared operational view should be clearer for readers outside the airport context, and suggested renaming cross-boundary information exchange to cross-organisational information exchange (I1, R). In the refined framework, D2.2 was renamed *Shared operational data layer*, because the original wording overlapped with later dimensions on visual representation and human interaction. D2.3 was retained but sharpened by referring to standardized identifiers, formats, and interfaces.

The interviews also showed that system integration is not only a technical issue. Schiphol participants explained that external data from airlines or handlers can be harder to integrate than internal Schiphol data because it may be commercially sensitive (I3, S1–S3). This was confirmed in the KLM interview, where A2 stated:

> “*We do not share all data with Schiphol, other than what we legally have to.*” (I4, A2)

Overall, D2 was retained but simplified. D2.1 was renamed *Integration of direct turnaround stakeholder data*, D2.2 was renamed *Shared operational data layer*, D2.3 was retained with sharper wording, and D2.4 was removed from the internal maturity indicators. The refined D2 indicators are shown in Table 6.4.

Table 6.4: Refined D2 system integration indicators after expert feedback

Code	Indicator	Indicator definition
D2 System integration		
D2.1	Integration of direct turnaround stakeholder data	Assesses whether data from direct turnaround stakeholders are connected to the digital twin environment.
D2.2	Shared operational data layer	Assesses whether operational data are integrated into one consistent data layer for the digital twin.
D2.3	Standardized identifiers and interfaces	Assesses whether connected systems use shared identifiers, formats, and interfaces to ensure that different sources refer to the same operational objects.

6.5. Feedback on D3 Predictive and analytical capability

The interviewees reviewed the literature-based D3 indicators presented in Appendix B. These indicators assessed whether the digital twin could detect emerging problems, estimate operational impacts, explore scenarios, use data-driven prediction, and support causal analysis. The literature-based table is not repeated here, because this section focuses on how the expert feedback led to the refined indicators.

Overall, D3 Predictive and analytical capability was confirmed as a relevant dimension. The dimension captures the step from monitoring the current turnaround situation to understanding what may happen next. C2 described D1 to D4 as a logical sequence, in which data infrastructure and system integration enable prediction, which then supports decision support and optimisation (I5, C2). Therefore, D3 was retained as a separate internal maturity dimension.

The main refinement concerned the boundary between D3 and D4. R noted that the two dimensions are closely related, but have a different focus:

> “*One is really looking at the data and saying: that might happen. The other one is really: what should we do about it?*” (I1, R)

This distinction was also supported in the Schiphol focus group. S1 explained that prediction concerns identifying what is likely to happen, while decision support moves towards selecting a response:

> “*Predictive says more: this is what is going to happen. Decision support goes more towards the solution.*” (I3, S1, translated)

Based on this feedback, D3 was defined more clearly as the analytical capability of the digital twin. D3 therefore focuses on prediction, early problem detection, impact estimation, and scenario analysis, while D4 focuses on using these insights to support or optimise operational decisions.

The interviews also confirmed the relevance of historical data and scenario analysis for predictive maturity. Schiphol participants explained that historical patterns can help users understand whether similar situations occurred before and what effects they had (I3, S1–S3). The KLM interview supported this point through the example of ALIGN, which can be used to simulate disruptions and assess their effects on schedules and delays, although it is not a real-time day-of-operation digital twin (I4, A1–A2). This supported keeping scenario analysis as a core D3 indicator.

Overall, D3 was retained, but the order and scope of the indicators were sharpened. *Data-driven prediction* was moved to the start of the dimension, because it forms the basis for the other analytical capabilities. *Early problem detection*, *Impact estimation*, and *Scenario analysis* were retained with clearer wording. *Causal analysis support* was removed as a separate indicator, because its purpose partly overlapped with early problem detection and impact estimation. The refined D3 indicators are shown in Table 6.5.

Table 6.5: Refined D3 predictive and analytical capability indicators after expert feedback

Code	Indicator	Indicator definition
D3 Predictive and analytical capability		
D3.1	Data-driven prediction	Assesses whether the digital twin uses live and historical data to predict future turnaround performance.
D3.2	Early problem detection	Assesses whether the digital twin identifies emerging turnaround problems before they fully materialise.
D3.3	Impact estimation	Assesses whether the digital twin estimates which parts of the operation are likely to be affected by disruptions, deviations, or changing conditions, and which are not.
D3.4	Scenario analysis	Assesses whether the digital twin explores what-if scenarios and shows their expected operational effects.

6.6. Feedback on D4 Decision support and optimisation

The interviewees reviewed the literature-based D4 indicators presented in Appendix B. These indicators assessed whether the digital twin could support operational response selection, resource adjustment, task and schedule adjustment, comparison of response options, and actionable recommendations. The literature-based table is not repeated here, because this section focuses on how the expert feedback led to the refined indicators.

Overall, D4 Decision support and optimisation was confirmed as a relevant dimension. The dimension captures the step from understanding what may happen to deciding what should be done in response. R explained this distinction between D3 and D4 as follows:

> “*One is really looking at the data and saying: that might happen. The other one is really: what should we do about it?*” (I1, R)

C2 also described D1 to D4 as a logical sequence, in which data infrastructure, system integration, and prediction form the basis for decision support (I5, C2). Therefore, D4 was retained as a separate internal maturity dimension.

The main refinement concerned the need to make D4 more action-oriented. In the Schiphol focus group, participants explained that decision support should not only show or predict a problem, but also help users decide what to do next. S1 described the difference between prediction and decision support clearly:

> “*Predictive says more: this is what is going to happen. Decision support goes more towards the solution.*” (I3, S1, translated)

The refined D4 indicators therefore focus on whether digital twin insights are translated into operational actions, such as selecting a response, reassigning resources, or adjusting task order, timing, or assignment.

The interviews also showed that decision support should not be confused with full automated execution. Schiphol participants emphasised that Wilbur currently supports decisions, but does not independently execute actions such as gate changes or bus assignments. S2 noted:

> "*There is support, so the system does not do it itself, but supports decision-making.*" (I3, S2, translated)

This distinction was supported by the KLM interview, which showed that Schiphol, airlines, and ground handlers each control different parts of the turnaround process (I4, A1–A2). As a result, D4 maturity depends not only on technical capability, but also on decision rights, workflow integration, accountability, and operational authority.

Overall, D4 was retained, but the indicators were made more concrete and less repetitive. D4.1 was renamed *Data-based response selection*. D4.2 and D4.3 were sharpened around resource reassignment and task or schedule adjustment. The original indicator on comparing response options was removed because it overlapped with impact estimation in D3. The original indicator on actionable recommendations was merged into *Workflow-ready action*, because a recommendation only becomes operationally useful when it can be confirmed, sent, logged, or routed through the workflow. The refined D4 indicators are shown in Table 6.6.

Table 6.6: Refined D4 decision support and optimisation indicators after expert feedback

Code	Indicator	Indicator definition
D4 Decision support and optimisation		
D4.1	Data-based response selection	Assesses whether operational responses are selected using digital twin insights rather than mainly user experience.
D4.2	Resource reassignment	Assesses whether the digital twin uses resource availability to identify or propose changes to turnaround-relevant resources.
D4.3	Task and schedule adjustment	Assesses whether the digital twin identifies or proposes changes to task order, timing, or assignment.
D4.4	Workflow-ready action	Assesses whether the digital twin provides a concrete action that can be confirmed, sent, logged, or routed through the workflow.

6.7. Feedback on D5 Operational representation

The interviewees reviewed the literature-based D5 indicators presented in Appendix B. These indicators assessed whether the digital twin clearly represents the current turnaround status, task dependencies, disruptions, and the level of operational detail needed to support understanding and decision-making. The literature-based table is not repeated here, because this section focuses on how the expert feedback led to the refined indicators.

Overall, D5 Operational representation was retained as a relevant dimension. It captures how operational data and analytical outputs are translated into a clear representation of the turnaround situation. R noted that D5 is close to D6, but still has a different focus: D5 concerns what is represented in the digital twin, while D6 concerns how users understand, use, and interact with that representation (I1, R).

The main refinement concerned the position of D5 in relation to D2 and D3. C2 questioned whether D5 partly overlapped with system integration and predictive capability, because status and disruption visibility depend on the underlying data and analysis. However, C2 also clarified the distinct role of D5:

> *"It is a visual translation."* (I5, C2, translated)

D5 was therefore retained, but framed more clearly as the visual and operational translation of data and analysis. It does not assess whether the data or prediction exist, but whether they are represented in a way that helps users understand the operational situation.

A second refinement concerned the overlap between D5.1 and D5.4. In the Schiphol focus group, participants noted that current status visibility and operational level of detail partly measure the same issue. If important detail is missing, the current situation is also less visible. At the same time, they emphasised that more detail is not always better:

> *"The challenge is: what do you show, to whom and at what moment, so that this person can make better decisions?"* (I3, S3, translated)

This led to a clearer focus on role-relevant information. D5.4 was therefore renamed *Role-relevant operational detail*, because maturity depends on showing the right information to the right user, rather than showing as much information as possible.

The interviews also suggested that the current turnaround situation should include relevant operational context, such as stand characteristics, docking availability, or other assets that influence what can happen during the turnaround (I3, S1–S2). This was incorporated into the refined wording of D5.1. Overall, D5 was retained, but the indicators were sharpened to avoid overlap and to emphasise visual clarity, interpretability, and role relevance. The refined D5 indicators are shown in Table 6.7.

Table 6.7: Refined D5 operational representation indicators after expert feedback

Code	Indicator	Indicator definition
D5 Operational representation		
D5.1	Current turnaround situation visibility	Assesses whether the current status of aircraft, tasks, resources, and relevant operational context is clearly represented.
D5.2	Task dependency visibility	Assesses whether relationships and dependencies between turnaround tasks are clearly shown.
D5.3	Disruption visibility	Assesses whether delays, disruptions, and abnormal events are clearly represented and easy to interpret.
D5.4	Role-relevant operational detail	Assesses whether users are shown the operational detail relevant to their role without information overload.

6.8. Feedback on D6 Human interaction and collaboration

The interviewees reviewed the literature-based D6 indicators presented in Appendix B. These indicators assessed whether the digital twin interface was understandable and usable for intended users, supported shared situational awareness, provided trustworthy information, and supported collaboration between turnaround actors. The literature-based table is not repeated here, because this section focuses on how the expert feedback led to the refined indicators.

Overall, D6 Human interaction and collaboration was retained as a relevant dimension. C2 explicitly distinguished D6 from the more technical and analytical dimensions, because it focuses on human–machine interaction:

> *"D6 is a logical separate dimension; it is really about human–machine interaction."* (I5, C2, translated)

This confirmed that digital twin maturity should not only assess what the system can technically do, but also whether users can understand, use, and trust it in daily turnaround work.

The main refinement concerned the relation between D5 and D6. R noted that D5 and D6 are closely related, because representation is part of how users interact with the digital twin (I1, R). The Schiphol focus group similarly connected understandability to how operational information is represented. D6 was therefore kept focused on the user side of the interface: whether the information is understandable, practically usable, and trusted by intended users.

A second refinement concerned trust. The Schiphol focus group suggested that trust should not be treated as part of governance, because governance concerns rules and compliance, while trust concerns whether users believe and rely on the system. S3 summarised this distinction as follows:

> *"Trust is about: do I trust the system and can I work with it?"* (I3, S3, translated)

Information trustworthiness was therefore retained in D6, while governance was treated separately as an external implementation constraint.

Finally, the interviews showed that shared situational awareness and collaboration support were difficult to assess as separate interface indicators. R noted that shared understanding should not be assessed by one user only, but by the group of people working with the digital twin (I1, R). The Schiphol focus group also clarified that having the same information position does not automatically mean that actors collaborate well (I3, S1–S3). C2 similarly found collaboration support difficult to interpret and linked it to external conditions such as data sharing, competition, and available resources (I5, C2). Therefore, these indicators were not retained as separate D6 indicators in the refined internal maturity model.

Overall, D6 was retained but simplified. The refined dimension focuses on three core user-facing indicators: *Understandability*, *Usability in practice*, and *Information trustworthiness*. The refined D6 indicators are shown in Table 6.8.

Table 6.8: Refined D6 human interaction and collaboration indicators after expert feedback

Code	Indicator	Indicator definition
D6 Human interaction and collaboration		
D6.1	Understandability	Assesses whether intended users can understand the interface and presented information without specialist data expertise.
D6.2	Usability in practice	Assesses whether intended users can use the digital twin effectively in daily turnaround work.
D6.3	Information trustworthiness	Assesses whether users consider the digital twin's information reliable enough for operational interpretation and decision-making.

6.9. Feedback on D7 Governance, interoperability, and trust

The interviewees reviewed the literature-based D7 indicators presented in Appendix B. These indicators assessed whether roles and responsibilities were clearly defined, data-sharing rules were formalised, privacy, access control, and security were embedded, the digital twin was aligned with operational accountability structures, and future interoperability was supported. The literature-based table is not repeated here, because this section focuses on how the expert feedback led to the refinement of D7.

The largest refinement concerned the status of D7 within the framework. The interviews showed that D7 did not assess the digital twin in the same way as D1 to D6. While D1 to D6 describe what the digital twin is technically and operationally capable of doing, D7 mainly describes whether the airport context allows these capabilities to be used in practice. C2 therefore supported separating governance from the internal maturity dimensions:

> *"I think it is very good to separate governance from those elements."* (I5, C2, translated)

D7 was therefore removed from the internal maturity model and reformulated as a set of external implementation constraints.

A second refinement concerned the combination of governance, interoperability, and trust. The Schiphol focus group showed that trust should not be treated as part of governance. Trust concerns whether users believe and rely on the information presented by the digital twin, while governance concerns rules, compliance, responsibilities, and decision rights (I3, S1–S3). Trust was therefore moved to D6 Human interaction and collaboration. Future interoperability was also no longer treated as a separate governance indicator, because it mainly relates to system integration and data-sharing conditions.

The interviews also showed why governance should be treated as a constraining or enabling layer. A digital twin may be technically capable of supporting higher maturity, while data ownership, privacy rules, competition, contracts, safety procedures, or decision rights prevent these capabilities from being used. The KLM interview clearly illustrated this limitation:

> *"We do not share all data with Schiphol, other than what we legally have to."* (I4, A2)

A2 further explained that some operational and passenger-related data are commercially sensitive, legally restricted, or protected by contractual obligations (I4, A1–A2). This means that low maturity scores may reflect contextual constraints rather than only technical immaturity.

Overall, D7 was removed as an internal maturity dimension and replaced by four external implementation constraints. These constraints assess what the airport environment allows the digital twin to do, rather than what the digital twin can technically do. The resulting external implementation constraints are shown in Table 6.9.

Table 6.9: External implementation constraints derived from the refinement and discussion of D7

External implementation constraint	**Core question**	**Description**
1. Data access, ownership, and privacy constraints	Can the required data be used?	Conditions that determine whether the digital twin is allowed and able to access, use, process, store, protect, and share the data it needs. This includes data ownership, data-sharing agreements, privacy and GDPR requirements, access rights, cybersecurity requirements, and restrictions on personal or sensitive data.
2. Commercial and competition constraints	Are stakeholders willing to share sensitive operational information?	Conditions that determine whether stakeholders are willing to share operational data that may be commercially sensitive. This includes data that could reveal performance, staffing, capacity, delays, costs, or operational weaknesses.

Continued on next page

Table 6.9: External implementation constraints continued

External implementation constraint	Core question	Description
3. Decision, accountability, and workflow constraints	Can digital twin outputs be used in the operational decision process?	Conditions that determine whether digital twin outputs can be used within existing decision rights, accountability structures, A-CDM agreements, approval procedures, fallback procedures, and operational workflows. These conditions define who may act, who must approve, who is responsible, and whether the twin fits the established way of working.
4. Safety constraints	Is the use of the digital twin allowed within safety boundaries?	Conditions that determine whether digital twin outputs or actions are allowed within aviation safety rules, airside procedures, and risk controls. These conditions may limit whether the twin can change tasks, gates, resources, priorities, or operational instructions, even if this is technically possible.

6.10. Feedback on thresholds

Table 6.10: Literature-based threshold logic repeated from Appendix B as the basis for expert feedback

Level	Label	Main capability	Key indicators and threshold
1	Digital model	Turnaround is represented digitally	No formal threshold; basic digital representation exists.
2	Digital shadow	Operational data automatically update the system	**D1.1** Automatic data capture ≥ 3; **D1.2** Timely data updating ≥ 3; **D1.3** Data completeness ≥ 2; **D2.1** Multi-stakeholder data integration ≥ 2
3	Monitoring twin	Shared monitoring across multiple actors	**D1.3** Data completeness ≥ 3; **D2.1** Multi-stakeholder data integration ≥ 3; **D2.2** Shared operational view ≥ 3; **D6.3** Shared situational awareness ≥ 3
4	Predictive twin	Forecasting and impact estimation	**D3.1** Early problem detection ≥ 3; **D3.2** Impact estimation ≥ 3; **D3.3** Data-driven prediction ≥ 3; **D3.4** Scenario analysis ≥ 3; **D3.5** Causal analysis support ≥ 3
5	Optimising twin	Decision support and dynamic optimisation	**D4.1** Response selection support ≥ 4; **D4.2** Resource adjustment support ≥ 4; **D4.3** Task and schedule adjustment support ≥ 4; **D4.4** Comparison of response options ≥ 4; **D4.5** Actionable recommendations ≥ 4

The interviewees also reviewed the preliminary threshold logic shown in Table 6.10. This logic connects maturity levels to minimum indicator scores and was intended to make the relationship between levels, dimensions, and indicators more explicit.

Overall, the threshold approach was considered useful as a structured proposal. R supported the idea of defining minimum scores for indicators that are necessary for reaching a maturity level:

> *"I think it makes quite a lot of sense to define that each indicator needs to reach a certain level."* (I1, R)

This makes the assessment more transparent, because weak performance on an essential indicator cannot be hidden by stronger scores elsewhere. C2 similarly noted that the relationship between levels, dimensions, and indicators needed to be made clearer (I5, C2).

At the same time, the interviews showed that the thresholds should be interpreted carefully. The Schiphol focus group explained that maturity can differ across parts of the turnaround process:

> *"On one aspect you may be at level 1, and on another aspect at level 4."* (I3, S1, translated)

Therefore, the thresholds should be used to identify gaps and improvement priorities, rather than to reduce the digital twin to one overall score. They should also be interpreted together with the external implementation constraints, because achievable maturity may be limited by data access, decision rights, safety rules, or other contextual conditions (I5, C2).

Based on this feedback, the threshold logic is retained as a structured assessment proposal, but not as a fully validated measurement instrument. In the next chapter, the refined threshold logic is discussed in more detail with the researcher before the final framework is presented.

7
Final Framework: Feedback on Refined Framework

This chapter forms the final part of answering the following sub-question:

SQ2 - Which maturity dimensions, levels, and assessment logic should a digital twin maturity assessment framework for airport turnaround operations contain?

The previous chapter presented the refined maturity assessment framework based on expert feedback. This chapter presents the final maturity assessment framework that is used in the demonstration and evaluation phase of this research. As shown in Figure 7.1, this chapter is the final step in the artefact development process.

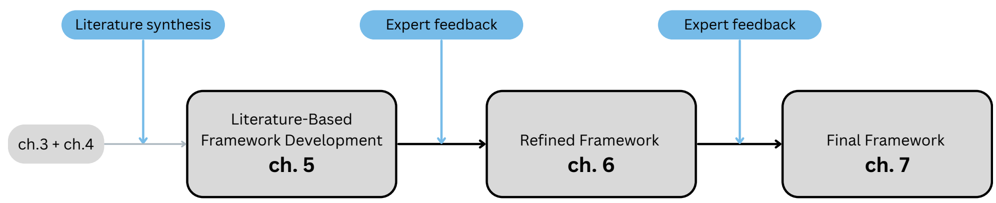


Figure 7.1: Development process of the maturity assessment framework repeated from Figure 5.1

Figure 7.1 is repeated to clarify that this chapter finalises the framework after the literature-based development and expert refinement steps. The follow-up interview was used to review the refined framework and to make the final adjustments before the framework was applied in the demonstration phase.

The methodological setup of the follow-up interview is described in Chapter 2. The interview materials and participant overview are provided in Appendix D. Therefore, this chapter focuses on how the follow-up feedback informed the final framework. The expert feedback is reported using the interview and participant code (I6, R). The chapter then presents the final maturity levels, the external implementation constraints, the final dimensions and indicators, and the scoring and threshold logic that together form the maturity assessment framework.

7.1. Feedback on the refined model

The follow-up interview confirmed that the refined framework was moving in a suitable direction. The researcher supported the separation between external implementation constraints and the internal maturity model, because the external constraints define what the airport context allows before the internal digital twin capabilities are assessed:

> *"I think it is fine to do it like that: really setting the boundaries, as you said, and then saying, within these possible boundaries, we are now applying the model."* (I6, R)

This feedback supported the final distinction between internal maturity and external implementation constraints. Internal maturity assesses what the digital twin is technically and operationally capable of, while the external constraints assess whether these capabilities can actually be used in practice.

The interview also led to the addition of a baseline level. The researcher noted that a level was missing for situations in which the digital twin is not yet present or not mature enough:

> *"The visualization makes sense. The only thing that just came to mind is: is there something like level zero?"* (I6, R)

Therefore, a baseline level was added to indicate when the minimum conditions for digital twin maturity are not yet met. Related to this, zero scores are shown explicitly instead of leaving cells empty, because this makes clear that an indicator has been assessed even when the capability is absent.

A further point concerned the order of assessment. The researcher suggested starting with the external implementation constraints, because this first defines the practical boundaries of the digital twin:

> *"My first thought would be to start with the external implementation constraints, because that really sets the boundaries of the digital twin."* (I6, R)

At the same time, the interview showed that the preferred order depends on the goal of the assessment. Starting internally may show the technical potential of the digital twin, while starting externally gives a more realistic view of what can be used within the current airport context. For the Schiphol demonstration, the external constraints were therefore assessed first.

Finally, the researcher emphasised that the framework should remain detailed enough to produce meaningful results. A simplified model would be easier to apply, but could hide important differences between dimensions and indicators:

> *"You can also make it too easy. If it becomes too easy, the result is not useful at all."* (I6, R)

This supported the decision to keep separate dimensions, indicators, scoring descriptions, and threshold logic. However, the scoring descriptions and thresholds are treated as a structured application proposal rather than as a fully validated measurement instrument.

Based on the follow-up feedback, the refined model was finalised into the framework presented below. The following sections present the final maturity levels, external implementation constraints, internal maturity dimensions and indicators, and threshold logic used in the demonstration phase.

7.2. Overview of the final maturity assessment framework

The final artefact of this research is a maturity assessment framework for airport turnaround digital twins. The framework does not only classify whether a digital twin exists, but assesses how mature its capabilities are and what limits its further development. It therefore combines two perspectives: internal maturity and external implementation constraints. Internal maturity concerns what the digital twin is technically and operationally capable of doing, while external implementation constraints concern what the airport context allows the digital twin to do in practice.

The framework consists of four connected components. Component 1 is the maturity level scale, which defines the development path of the digital twin. Component 2 consists of the external implementation constraints, which define what level of maturity can realistically be used in the airport context. Component 3 consists of the internal maturity dimensions and indicators, which define what capabilities are assessed. Component 4 consists of the threshold logic, which explains how the indicator scores are translated into an internal maturity level. Together, these components produce a maturity profile and identify capability gaps.

Figure 7.2 provides an abstract overview of how these components fit together. The figure shows that the assessment first distinguishes between external constraints and internal maturity. The external constraint level indicates what is practically achievable in the implementation context, while the internal maturity level indicates what the digital twin itself is capable of delivering. The gap between these two levels helps identify development priorities.

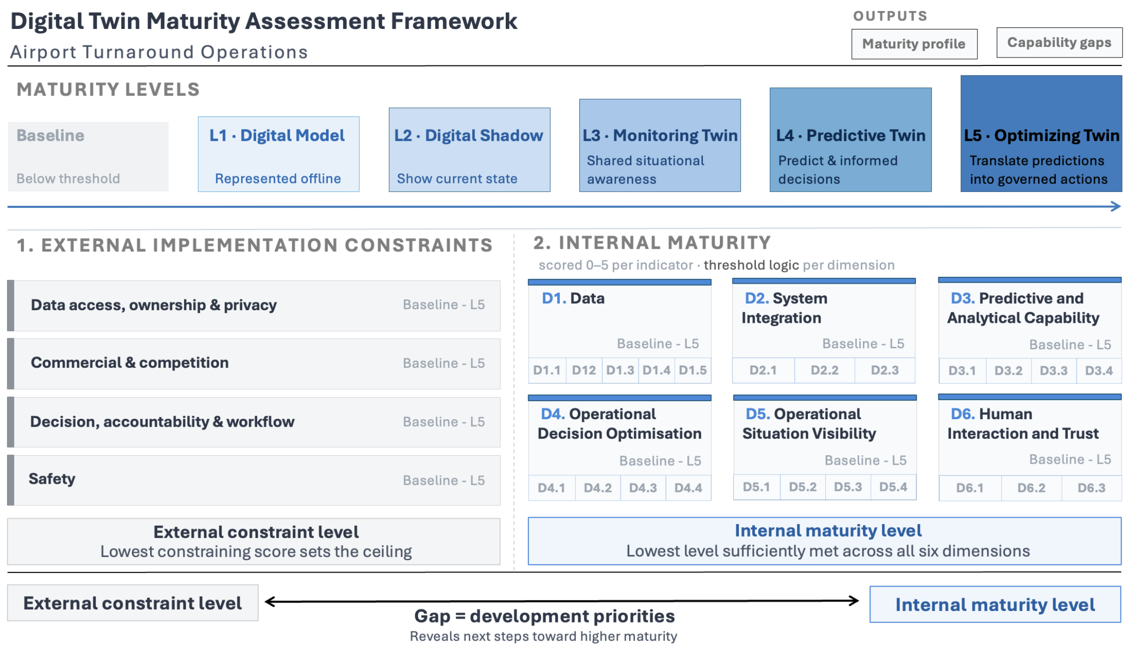


Figure 7.2: Abstract overview of the final digital twin maturity assessment framework for airport turnaround operations

The main analytical value of the framework is that it separates three questions that are often combined in practice: what the digital twin can technically do, what the airport context allows it to do, and what needs to improve next. This makes the assessment useful not only for assigning a maturity level, but also for discussing development priorities.

7.3. Application of the final maturity assessment framework

This section explains how the final maturity assessment framework can be applied in practice. The framework combines maturity levels, external implementation constraints, internal maturity dimensions, indicators, and scoring logic.

The assessment consists of two steps. First, the external implementation constraints are assessed. These constraints indicate what level of maturity can realistically be supported in the current airport context. They cover data access, ownership and privacy; commercial and competition constraints; decision-making, accountability and workflow; and safety. The lowest constraining score determines the overall external constraint level.

Second, the internal maturity of the digital twin is assessed. Internal maturity concerns what the digital twin itself is capable of delivering across six dimensions: data, system integration, predictive and analytical capability, operational decision optimisation, operational situation visibility, and human interaction and trust. Each dimension consists of indicators scored from 0 to 5. This scale broadly follows the maturity level scale: 0 indicates that the capability is below the maturity threshold, while scores 1 to 5 indicate increasing maturity from basic digital representation towards monitoring, prediction, and optimisation-oriented decision support. The exact scoring descriptions differ per indicator, and the scores are combined through threshold logic to determine the internal maturity level.

The framework produces two outputs: a maturity profile and capability gaps. The maturity profile shows both the internal maturity level and the external constraint level. The gap between them indicates where development is needed. If internal maturity is lower, the digital twin capabilities need further development. If the external constraint level is lower, the main limitation lies in the implementation context, such as data access, decision rights, workflows, or safety boundaries.

7.4. Component 1: Maturity level scale

The maturity level scale defines the development path of an airport turnaround digital twin, from basic digital representation to current-state visibility, shared monitoring, prediction, and optimisation-oriented decision support. The baseline level was added after the follow-up interview to show when the minimum conditions for digital twin maturity are not yet met.

The levels should be read as a progression of operational roles rather than as separate labels. Lower levels mainly represent or show the turnaround process, while higher levels support anticipation and structured operational action. The final maturity level scale is presented in Table 7.1.

Table 7.1: Final framework component 1: maturity level scale for airport turnaround digital twins

Level	**Name**	**Clear role of the digital twin**	**Description**
Baseline	Below Maturity Threshold	Capability not yet sufficient	The required conditions for the first applicable maturity level are not yet met.
Level 1	Digital Model	Represent offline	A static or non-real-time digital version of the turnaround process. It is mainly used for documentation, offline analysis, or internal understanding. It does not reflect or influence live operations.
Level 2	Digital Shadow	Show current state	A real-time or near-real-time overview of the turnaround operation. The system shows the current operational state, but this view is relatively simple and not yet connected across different turnaround stakeholders. Information mainly flows one way: from the physical operation to the digital system.
Level 3	Monitoring Twin	Create shared situational awareness	A digital twin that combines live information from multiple turnaround actors into one shared view. It shows what is happening now and makes the actions and progress of different actors visible, such as aircraft status, task progress, delays, and resource availability. Operational decisions and actions remain with the existing actors and workflows.
Level 4	Predictive Twin	Predict and inform decisions	A digital twin that helps users understand what is likely to happen next. It uses live and historical data to predict future states, identify risks, and estimate possible consequences for the operation. It informs decision-making, but does not yet translate predictions into structured operational actions.
Level 5	Optimizing Twin	Translate predictions into action	A digital twin that uses predictive insights to identify and prioritise operational decisions, such as task resequencing, resource reassignment, or schedule adjustment. It may propose or initiate decisions through connected systems, while physical execution remains with the responsible actors within defined decision, workflow, accountability, and safety boundaries.

7.5. Component 2: External implementation constraints

External implementation constraints define what level of digital twin maturity can realistically be used in the airport context. These constraints are included because a digital twin may be technically capable of prediction, optimisation, or recommendation, while the airport environment may still limit the use of these capabilities. Such limitations can come from data access, privacy requirements, commercial sensitivity, decision rights, workflows, accountability arrangements, or safety boundaries.

This component therefore adds an important analytical layer to the framework. It prevents the assessment from treating maturity as a purely technical issue and helps explain why a digital twin may not be used to its full technical potential. Table 7.2 defines the four external implementation constraints. Table 7.3 then translates these constraints into maturity-level requirements, showing what contextual conditions must be in place for each level to be practically achievable.

Table 7.2: Final framework component 2a: external implementation constraint definitions

External implementation constraint	Core question	Description
1. Data access, ownership, and privacy constraints	Can the required data be used?	Conditions that determine whether the digital twin is allowed and able to access, use, process, store, protect, and share the data it needs. This includes data ownership, data-sharing agreements, privacy and GDPR requirements, access rights, cybersecurity requirements, and restrictions on personal or sensitive data.
2. Commercial and competition constraints	Are stakeholders willing to share sensitive operational information?	Conditions that determine whether stakeholders are willing to share operational data that may be commercially sensitive. This includes data that could reveal performance, staffing, capacity, delays, costs, or operational weaknesses.
3. Decision, accountability, and workflow constraints	Can digital twin outputs be used in the operational decision process?	Conditions that determine whether digital twin outputs can be used within existing decision rights, accountability structures, A-CDM agreements, approval procedures, fallback procedures, and operational workflows. These conditions define who may act, who must approve, who is responsible, and whether the twin fits the established way of working.
4. Safety constraints	Is the use of the digital twin allowed within safety boundaries?	Conditions that determine whether digital twin outputs or actions are allowed within aviation safety rules, airside procedures, and risk controls. These conditions may limit whether the twin can change tasks, gates, resources, priorities, or operational instructions, even if this is technically possible.

Table 7.3: Final framework component 2b: external implementation constraint requirements by maturity level

Maturity level	**Description**	**External implementation constraint**			
		Data access, ownership and privacy	**Commercial and competition**	**Decision, accountability and workflow**	**Safety**
Baseline – Below Maturity Threshold	The required conditions for the first applicable maturity level are not yet met.	Required data conditions are not yet sufficient.	Required stakeholder conditions are not yet sufficient.	Required decision and workflow conditions are not yet sufficient.	Required safety conditions are not yet sufficient.
Level 1 – Digital Model	A static or non-real-time digital version of the turnaround process. It is mainly used for documentation, offline analysis, or internal understanding. It does not reflect or influence live operations.	Basic process data can be used for offline modelling, documentation, or internal analysis.	External stakeholder data can be avoided, anonymised, aggregated, or used only in a limited way.	The model can be created and interpreted without affecting live operational decisions.	The model can be used without affecting live operations or safety-critical processes.
Level 2 – Digital Shadow	A real-time or near-real-time overview of the turnaround operation. The system shows the current operational state, but this view is relatively simple and not yet connected across different turnaround stakeholders. Information mainly flows one way: from the physical operation to the digital system.	Live or near-real-time operational data can be accessed, protected, and displayed.	Operational data can be shown to authorised users without exposing commercially sensitive information.	The system can be used as an operational overview only, without recommendations, triggers, or workflow changes.	Live information can be displayed without triggering safety-critical actions or creating unsafe interpretation.
Level 3 – Monitoring Twin	A digital twin that combines live information from multiple turnaround actors into one shared view. It shows what is happening now and makes the actions and progress of different actors visible, such as aircraft status, task progress, delays, and resource availability. Operational decisions and actions remain with the existing actors and workflows.	Live data from multiple turnaround actors can be shared and combined into one shared operational view.	Stakeholders can share the monitoring data needed to show status, progress, delays, and resource availability, with sensitive data filtered or role-based where needed.	The shared view can be used for situational awareness and coordination, while operational decisions and actions remain within existing roles and workflows.	Safety-relevant information can be shown accurately and clearly, aligned with airside procedures.
Level 4 – Predictive Twin	A digital twin that helps users understand what is likely to happen next. It uses live and historical data to predict future states, identify risks, and estimate possible consequences for the operation. It informs decision-making, but does not yet translate predictions into structured operational actions.	Live and historical data can be used for prediction under clear rules for ownership, privacy, access, and data quality.	Stakeholders can share enough operational and historical data for meaningful prediction, while commercially sensitive patterns remain protected.	Predictions with clear rules on interpretation, accountability, escalation, and traceability. The twin does not yet translate predictions into actions.	Predictions can be interpreted within safety rules and operational limits.
Level 5 – Optimising Twin	A digital twin that uses predictive insights to identify and prioritise operational decisions, such as task resequencing, resource reassignment, or schedule adjustment. It may propose decisions through connected systems, while physical execution remains with the responsible actors within defined decision, workflow, accountability, and safety boundaries.	High-quality live and historical data can be used to translate predictions into structured operational actions.	Stakeholders allow governed use of operationally or commercially sensitive data for actions that may affect priorities, resources, or task sequences.	The twin's recommendations or actions can enter the operational decision process. Decision rights, human approval, override, fallback, accountability, and workflow integration are clearly defined.	Optimisation actions can be safely recommended or supported within airside safety boundaries, with validation, human override, fallback procedures, and rollback options.

7.6. Component 3: Internal maturity dimensions and indicators

The internal maturity model assesses the maturity of the digital twin itself. It consists of six dimensions: data, system integration, predictive and analytical capability, operational decision optimisation, operational situation visibility, and human interaction and trust. These dimensions show that maturity is not a single capability: a digital twin may be strong in data and visibility, while still being limited in prediction, scenario analysis, or workflow-ready decision support.

Table 7.4 provides a compact overview of the final dimensions and indicators. The first dimensions capture the data and integration foundation, the middle dimensions capture prediction and decision support, and the final dimensions capture operational visibility, usability, and trust.

The detailed 0–5 scoring descriptions are provided in Appendix E. This section therefore focuses on the structure and analytical logic of the framework, while the appendix provides the scoring detail needed for application.

Table 7.4: Final framework component 3: Internal maturity dimensions and indicators

Dimension	Code	Indicator	Indicator definition
D1 Data	D1.1	Automatic data capture	Assesses whether relevant turnaround data are captured automatically from operational sources.
	D1.2	Timely data availability	Assesses whether captured data become available quickly enough for the intended operational use.
	D1.3	Data completeness	Assesses whether the digital twin contains the core turnaround data needed for representation, prediction, and decision support.
	D1.4	Data quality management	Assesses whether data quality issues are systematically identified and addressed.
	D1.5	Historical operational data	Assesses whether historical turnaround data are stored and used for learning, prediction, and improvement.
D2 System integration	D2.1	Integration of direct turnaround stakeholder data	Assesses whether data from direct turnaround stakeholders are connected to the digital twin environment.
	D2.2	Shared operational data layer	Assesses whether operational data are integrated into one consistent data layer for the digital twin.
	D2.3	Standardised identifiers and interfaces	Assesses whether connected systems use shared identifiers, formats, and interfaces to ensure that different sources refer to the same operational objects.
D3 Predictive and analytical capability	D3.1	Data-driven prediction	Assesses whether the digital twin uses live and historical data to predict future turnaround performance.
	D3.2	Early problem detection	Assesses whether the digital twin identifies emerging turnaround problems before they fully materialise.
	D3.3	Impact estimation	Assesses whether the digital twin estimates which parts of the operation are likely to be affected by disruptions, deviations, or changing conditions, and which are not.

Continued on next page

Table 7.4: Final internal maturity dimensions and indicators continued

Dimension	Code	Indicator	Indicator definition
	D3.4	Scenario analysis	Assesses whether the digital twin explores what-if scenarios and shows their expected operational effects.
D4 Operational decision optimisation	D4.1	Data-based response selection	Assesses whether operational responses are selected using digital twin insights rather than mainly user experience.
	D4.2	Resource reassignment	Assesses whether the digital twin uses resource availability to identify or propose changes to turnaround-relevant resources.
	D4.3	Task and schedule adjustment	Assesses whether the digital twin identifies or proposes changes to task order, timing, or assignment.
	D4.4	Workflow-ready action	Assesses whether the digital twin provides a concrete action that can be confirmed, sent, logged, or routed through the workflow.
D5 Operational situation visibility	D5.1	Current turnaround situation visibility	Assesses whether the current status of aircraft, tasks, and resources is clearly represented.
	D5.2	Task dependency visibility	Assesses whether relationships and dependencies between turnaround tasks are clearly shown.
	D5.3	Disruption visibility	Assesses whether delays, disruptions, and abnormal events are clearly represented and easy to interpret.
	D5.4	Role-relevant operational detail	Assesses whether users are shown the operational detail relevant to their role without information overload.
D6 Human interaction and trust	D6.1	Understandability	Assesses whether intended users can understand the interface and presented information without specialist data expertise.
	D6.2	Usability in practice	Assesses whether intended users can use the digital twin effectively in daily turnaround work.
	D6.3	Information trustworthiness	Assesses whether users consider the digital twin's information reliable enough for operational interpretation and decision-making.

7.7. Component 4: Threshold logic

The threshold logic explains how the internal indicator scores are translated into an internal maturity level. Instead of relying only on averages, the framework defines minimum requirements for each level. This prevents strong performance in one dimension from compensating for missing capabilities in another essential dimension.

The logic treats maturity as a staged development process: higher levels can only be reached when the required foundational capabilities from earlier levels are sufficiently present. The 0–5 indicator scores broadly correspond to the maturity level scale, so a higher score means that a specific capability is more mature. However, the threshold logic does not require every indicator to reach the same score as the overall maturity level. Some indicators define a level, while others only support it.

For example, Level 2 mainly means that the digital twin can show the current operational state using operational data. This requires some connection to operational data sources, but not yet full integration across all relevant stakeholder systems. A Level 2 digital twin can therefore still rely on limited data feeds, partial integration, or one-way data flows. More advanced integration becomes necessary for Level 3, where the digital twin should support shared monitoring across multiple turnaround actors.

A 0–5 scoring scale is still used because it keeps the indicator scores connected to the overall maturity logic. Each indicator can be assessed along the same development path as the maturity levels: from absent or not assessable, to basic representation, current-state visibility, shared monitoring, prediction, and optimisation-oriented decision support. The threshold logic then adds flexibility by defining which indicators must reach which score for each overall maturity level.

Not every dimension is required at every maturity level. Lower levels focus mainly on data, visibility, and usability, while predictive and optimisation capabilities appear only once a sufficient data, integration, and monitoring foundation is present. Empty cells therefore mean that the dimension is not yet required for that level. The final internal maturity level requirements are presented in Table 7.5.

Table 7.5: Final framework component 4: internal maturity level threshold logic

D1.1	D1.2	D1.3	D1.4	D1.5	D2.1	D2.2	D2.3	D3.1	D3.2	D3.3	D3.4	D4.1	D4.2	D4.3	D4.4	D5.1	D5.2	D5.3	D5.4	D6.1	D6.2	D6.3

Internal maturity level

Maturity level	**Minimum dimension requirements**					
	Data	**System integration**	**Predictive and analytical capability**	**Operational decision optimisation**	**Operational situation visibility**	**Human interaction and trust**
Baseline – Below Maturity Threshold	D1 $<$ Level 1	D2 $<$ Level 2	D3 $<$ Level 3	D4 $<$ Level 3	D5 $<$ Level 1	D6 $<$ Level 1
Level 1 – Digital Model	D1.1 ≥ 1; D1.3 ≥ 1; D1.4 ≥ 1				D5.1 ≥ 1; D5.2 ≥ 1; D5.4 ≥ 1	D6.1 ≥ 1; D6.2 ≥ 1
Level 2 – Digital Shadow	D1.1 ≥ 2; D1.2 ≥ 2; D1.3 ≥ 2; D1.4 ≥ 2	D2.1 ≥ 1; D2.2 ≥ 1; D2.3 ≥ 1			D5.1 ≥ 2; D5.3 ≥ 1; D5.4 ≥ 2	D6.1 ≥ 2; D6.2 ≥ 2; D6.3 ≥ 2
Level 3 – Monitoring Twin	D1.1 ≥ 3; D1.2 ≥ 3; D1.3 ≥ 3; D1.4 ≥ 3	D2.1 ≥ 3; D2.2 ≥ 3; D2.3 ≥ 3	D3.2 ≥ 2	D4.1 ≥ 2; D4.2 ≥ 2; D4.3 ≥ 2; D4.4 ≥ 1	D5.1 ≥ 3; D5.2 ≥ 3; D5.3 ≥ 3; D5.4 ≥ 3	D6.1 ≥ 3; D6.2 ≥ 3; D6.3 ≥ 3
Level 4 – Predictive Twin	D1.1 ≥ 4; D1.2 ≥ 4; D1.3 ≥ 4; D1.4 ≥ 4; D1.5 ≥ 4	D2.1 ≥ 4; D2.2 ≥ 4; D2.3 ≥ 4	D3.1 ≥ 4; D3.2 ≥ 4; D3.3 ≥ 4; D3.4 ≥ 3	D4.1 ≥ 3; D4.2 ≥ 3; D4.3 ≥ 3; D4.4 ≥ 2	D5.1 ≥ 4; D5.3 ≥ 4; D5.4 ≥ 4	D6.1 ≥ 4; D6.2 ≥ 4; D6.3 ≥ 4
Level 5 – Optimising Twin	D1.1 $= 5$; D1.2 $= 5$; D1.3 $= 5$; D1.4 $= 5$; D1.5 $= 5$	D2.1 $= 5$; D2.2 $= 5$; D2.3 $= 5$	D3.1 $= 5$; D3.2 $= 5$; D3.3 $= 5$; D3.4 ≥ 4	D4.1 $= 5$; D4.2 ≥ 4; D4.3 ≥ 4; D4.4 $= 5$	D5.1 $= 5$; D5.2 ≥ 4; D5.3 $= 5$; D5.4 $= 5$	D6.1 $= 5$; D6.2 $= 5$; D6.3 $= 5$

Part D

Demonstrate Artefact

8 Demonstration at Schiphol

This chapter answers the following sub-question:

SQ3 - To what extent can the proposed maturity assessment framework be meaningfully applied in a real airport turnaround context?

This chapter demonstrates whether and how the maturity assessment framework can be applied in practice. The framework is applied to the Schiphol turnaround context to assess the maturity of a turnaround-related digital twin application and to identify the main constraints and development steps. The purpose of this chapter is therefore not to evaluate the framework itself, but to show how the framework can be used in a real airport turnaround setting.

8.1. Schiphol turnaround digital twin context

The framework was demonstrated in the Schiphol turnaround context. The assessed digital twin application is related to Wilbur and can be understood as a dashboard-based digital twin environment for monitoring and supporting airport operations. Within Wilbur, the turnaround process is represented as a specific tab in the dashboard. The aim of the assessment was to determine what this turnaround-related digital twin capability is currently able to deliver in relation to operational visibility, prediction, decision support, and turnaround decision-making.

In this demonstration, the focus is not on assessing Schiphol as an airport in general, but on assessing the maturity of the turnaround-related digital twin capabilities. This means that the assessment examines how data are captured, integrated, represented, analysed, and used to support the turnaround process. The demonstration therefore provides insight into both the internal capabilities of the digital twin and the external implementation constraints that influence whether these capabilities can be used in practice.

8.2. Application procedure

Chapter 7 explained the general application logic of the maturity assessment framework. In this chapter, that logic is applied to the Schiphol case. The demonstration followed two main steps. First, the external implementation constraints were assessed through a group discussion. This was done to create a shared understanding of the operational, organisational, data-related, and safety-related environment in which the digital twin has to function.

Second, the internal maturity dimensions were assessed. Participants first scored the internal maturity indicators individually. Afterwards, the individual scores were discussed in the group to arrive at one shared score per indicator. This procedure was used to prevent the assessment from being based only on one dominant interpretation and to support a more detailed discussion of the reasoning behind the scores.

The full scoring tables and supporting notes are provided in Appendix F. The main text focuses on the overall scoring pattern, the interpretation of the results, and the development steps that follow from the demonstration.

8.3. Demonstration results

The demonstration resulted in two connected outcomes: an external constraint level and an internal maturity level. The external constraint level shows what level of maturity is practically feasible within the Schiphol implementation context. The internal maturity level shows what the digital twin itself is currently capable of delivering. The detailed scores are included in Appendix F. Figures 8.1 and 8.2 provide an overview of the demonstration scores.

8.3.1. External implementation constraints

The external implementation constraints show that the practical maturity of the Schiphol turnaround digital twin is not limited equally across all conditions. Data access, ownership, and privacy were assessed as relatively mature. Participants indicated that role-based access and data protection mechanisms are already in place. This means that users can access information that is relevant to their role, while sensitive information can be filtered. However, the system does not own all source data itself and not all desired data are available or complete. Therefore, this constraint was not assessed at the highest level.

Commercial and competition constraints were assessed between monitoring and predictive maturity. Participants explained that much operational information is already shared, such as arrival times, departure times, and other contextual data. At the same time, not all commercially sensitive information is available to all users. Some information is filtered by role or organisation. This means that the constraint is mature enough to support shared monitoring, but not yet fully mature for all predictive or optimisation purposes.

Decision, accountability, and workflow constraints were also assessed between monitoring and predictive maturity. The system can support operational interpretation and coordination, but its outputs are not yet fully embedded in formal decision rights, escalation procedures, traceability rules, or operational workflows. Some parts of the operation, such as gate planning, are more advanced, while turnaround handling itself is closer to monitoring maturity.

Safety was the lowest-scoring external constraint. Participants indicated that Wilbur can display live operational information, but that the digital twin is not yet sufficiently aligned with airside safety procedures to rely on it for safety-relevant operational decisions. The system still requires human judgement and additional checks, especially when advice could influence gate use, aircraft positioning, resource allocation, or other operational decisions. This means that safety is the main external factor limiting the practical maturity of the system.

8.3.2. Internal maturity dimensions

The internal maturity scores show a mixed maturity profile. The strongest scores were found in timely data availability (D1.2), system integration, especially the shared operational data layer (D2.2) and standardised identifiers and interfaces (D2.3), data-driven prediction (D3.1), current turnaround situation visibility (D5.1), understandability (D6.1), and usability in practice (D6.2). This indicates that the digital twin already supports a live or near-real-time operational view and that users can generally interpret and use the system in practice.

However, several indicators received lower scores and therefore limit the overall internal maturity level. Data completeness (D1.3) and data quality management (D1.4) were assessed as relatively low. Participants mentioned missing ramp coverage, missing asset status, limited information on towed aircraft, and incomplete integration with handler resources and planning. These limitations reduce the extent to which the digital twin can represent the full turnaround process as one integrated operational system.

The predictive and analytical capability dimension (D3) shows that the digital twin already has some predictive functionality. For example, Deep Turnaround and TOBT-related advice indicate that the system can use operational data to support data-driven prediction (D3.1). However, scenario analysis (D3.4) remains limited, because the system does not yet compare alternative what-if scenarios or systematically estimate the effects of different operational responses.

The lowest scores were found in the operational decision optimisation dimension (D4). Although users consult the digital twin when interpreting the operation, operational responses are still mainly selected through professional experience and existing procedures. The system does not yet structurally support resource reassignment (D4.2), task and schedule adjustment (D4.3), or workflow-ready actions (D4.4). This shows that the current system supports operational awareness more strongly than operational optimisation.

Operational situation visibility (D5) also shows a mixed pattern. The current turnaround situation (D5.1) and key disruptions (D5.3) are visible for important parts of the operation. However, task dependency visibility (D5.2) is still limited. This means that the system shows several elements of the turnaround situation, but does not yet clearly show how tasks depend on each other or how disruptions create knock-on effects across the turnaround process.

8.3.3. Comparison between external constraints and internal maturity

The comparison between the external constraints and the internal maturity scores shows that the Schiphol turnaround digital twin has developed beyond a basic digital model. Internally, the system already supports live operational visibility, shared interpretation, and some predictive insight. Several indicators score at monitoring or predictive levels, especially in relation to data availability, system integration, prediction, visibility, and usability.

At the same time, the internal scores show that the system does not yet meet the requirements for a full Monitoring Twin across the complete turnaround scope. This is mainly caused by lower scores on several specific indicators: data completeness (D1.3), data quality management (D1.4), resource reassignment (D4.2), task and schedule adjustment (D4.3), workflow-ready action (D4.4), and task dependency visibility (D5.2). These results show that the system can support operational awareness, but does not yet consistently connect information, predictions, and operational actions.

The external constraints further limit the practical maturity. Although data access and several decision-related conditions are relatively mature, the safety constraint remains at level 2. This means that even when the system has more advanced technical or predictive capabilities in specific areas, these capabilities cannot yet be fully used for operational recommendations or optimisation without additional safety alignment, validation, and human checks.

8.3.4. Overall maturity level

Based on the demonstration, the overall maturity of the Schiphol turnaround digital twin is interpreted as **Level 2: Digital Shadow**. The system provides a live or near-real-time operational view and supports users in understanding the current turnaround situation. However, because the system does not yet consistently combine all required data, show task dependencies, support workflow-ready actions, or meet the external safety conditions for higher maturity, it cannot yet be classified as a full Monitoring Twin across the selected turnaround scope.

This result does not mean that all parts of the system are limited to level 2. Several individual capabilities are more advanced, especially in relation to data availability, system integration, prediction, and usability. However, the threshold logic of the framework prevents higher scores in some indicators from compensating for missing minimum conditions in other indicators. The final maturity level therefore reflects the highest level that is sufficiently supported across the internal maturity dimensions and external implementation constraints.

8.4. Conclusion of the demonstration

The Schiphol demonstration shows that the maturity assessment framework can be applied in a real airport turnaround context. The framework helped structure the assessment by separating internal digital twin capabilities from external implementation constraints. This distinction proved useful because some limitations were not caused by the technical system itself, but by the operational, organisational, data-sharing, workflow, and safety context in which the digital twin has to function.

The demonstration also showed that the framework can identify specific development priorities. For Schiphol, the main development areas are improving data completeness and data quality management, representing task dependencies and knock-on effects more explicitly, developing scenario analysis, and translating predictive insights into workflow-ready operational actions. At the same time, further development towards higher maturity levels requires stronger alignment with decision rights, accountability structures, workflow integration, and safety validation.

Overall, the demonstration confirms that the framework is useful not only for assigning a maturity level, but also for making the maturity profile explicit. It shows where the current digital twin is already relatively mature, where capability gaps remain, and which external constraints limit further development. The Schiphol turnaround digital twin is therefore assessed as Level 2, while containing several capabilities that provide a foundation for further development towards monitoring and predictive maturity.

Table 8.1: Schiphol demonstration scores mapped onto the external implementation constraint requirements

Maturity level	Description	External implementation constraint			
		Data access, ownership and privacy	**Commercial and competition**	**Decision, accountability and workflow**	**Safety**
Baseline – Below Maturity Threshold	The required conditions for the first applicable maturity level are not yet met.	Required data conditions are not yet sufficient.	Required stakeholder conditions are not yet sufficient.	Required decision and workflow conditions are not yet sufficient.	Required safety conditions are not yet sufficient.
Level 1 – Digital Model	A static or non-real-time digital version of the turnaround process. It is mainly used for documentation, offline analysis, or internal understanding. It does not reflect or influence live operations.	Basic process data can be used for offline modelling, documentation, or internal analysis.	External stakeholder data can be avoided, anonymised, aggregated, or used only in a limited way.	The model can be created and interpreted without affecting live operational decisions.	The model can be used without affecting live operations or safety-critical processes.
Level 2 – Digital Shadow	A real-time or near-real-time overview of the turnaround operation. The system shows the current operational state, but this view is relatively simple and not yet connected across different turnaround stakeholders. Information mainly flows one way: from the physical operation to the digital system.	Live or near-real-time operational data can be accessed, protected, and displayed.	Operational data can be shown to authorised users without exposing commercially sensitive information.	The system can be used as an operational overview only, without recommendations, triggers, or workflow changes.	Live information can be displayed without triggering safety-critical actions or creating unsafe interpretation.
Level 3 – Monitoring Twin	A digital twin that combines live information from multiple turnaround actors into one shared view. It shows what is happening now and makes the actions and progress of different actors visible, such as aircraft status, task progress, delays, and resource availability. Operational decisions and actions remain with the existing actors and workflows.	Live data from multiple turnaround actors can be shared and combined into one shared operational view.	Stakeholders can share the monitoring data needed to show status, progress, delays, and resource availability, with sensitive data filtered or role-based where needed.	The shared view can be used for situational awareness and coordination, while operational decisions and actions remain within existing roles and workflows.	Safety-relevant information can be shown accurately and clearly, aligned with airside procedures.
Level 4 – Predictive Twin	A digital twin that helps users understand what is likely to happen next. It uses live and historical data to predict future states, identify risks, and estimate possible consequences for the operation. It informs decision-making, but does not yet translate predictions into structured operational actions.	Live and historical data can be used for prediction under clear rules for ownership, privacy, access, and data quality.	Stakeholders can share enough operational and historical data for meaningful prediction, while commercially sensitive patterns remain protected.	Predictions with clear rules on interpretation, accountability, escalation, and traceability. The twin does not yet translate predictions into actions.	Predictions can be interpreted within safety rules and operational limits.
Level 5 – Optimising Twin	A digital twin that uses predictive insights to identify and prioritise operational decisions, such as task resequencing, resource reassignment, or schedule adjustment. It may propose decisions through connected systems, while physical execution remains with the responsible actors within defined decision, workflow, accountability, and safety boundaries.	High-quality live and historical data can be used to translate predictions into structured operational actions.	Stakeholders allow governed use of operationally or commercially sensitive data for actions that may affect priorities, resources, or task sequences.	The twin's recommendations or actions can enter the operational decision process. Decision rights, human approval, override, fallback, accountability, and workflow integration are clearly defined.	Optimisation actions can be safely recommended or supported within airside safety boundaries, with validation, human override, fallback procedures, and rollback options.

Table 8.2: Schiphol demonstration scores mapped onto the internal maturity level requirements

D1.1	D1.2	D1.3	D1.4	D1.5	D2.1	D2.2	D2.3	D3.1	D3.2	D3.3	D3.4	D4.1	D4.2	D4.3	D4.4	D5.1	D5.2	D5.3	D5.4	D6.1	D6.2	D6.3
4	5	2	2	3	3	4	4	4	3	2	1	2	1	1	2	3	1	3	2	4	4	3

Internal maturity level

Maturity level	Minimum dimension requirements					
	Data	**System integration**	**Predictive and analytical capability**	**Operational decision optimisation**	**Operational situation visibility**	**Human interaction and trust**
Baseline – Below Maturity Threshold	D1 $<$ Level 1	D2 $<$ Level 2	D3 $<$ Level 3	D4 $<$ Level 3	D5 $<$ Level 1	D6 $<$ Level 1
Level 1 – Digital Model	$D1.1 \geq 1$; $D1.3 \geq 1$; $D1.4 \geq 1$				$D5.1 \geq 1$; $D5.2 \geq 1$; $D5.4 \geq 1$	$D6.1 \geq 1$; $D6.2 \geq 1$
Level 2 – Digital Shadow	$D1.1 \geq 2$; $D1.2 \geq 2$; $D1.3 \geq 2$; $D1.4 \geq 2$	$D2.1 \geq 1$; $D2.2 \geq 1$; $D2.3 \geq 1$			$D5.1 \geq 2$; $D5.3 \geq 1$; $D5.4 \geq 2$	$D6.1 \geq 2$; $D6.2 \geq 2$; $D6.3 \geq 2$
Level 3 – Monitoring Twin	$D1.1 \geq 3$; $D1.2 \geq 3$; $D1.3 \geq 3$; $D1.4 \geq 3$	$D2.1 \geq 3$; $D2.2 \geq 3$; $D2.3 \geq 3$	$D3.2 \geq 2$	$D4.1 \geq 2$; $D4.2 \geq 2$; $D4.3 \geq 2$; $D4.4 \geq 1$	$D5.1 \geq 3$; $D5.2 \geq 3$; $D5.3 \geq 3$; $D5.4 \geq 3$	$D6.1 \geq 3$; $D6.2 \geq 3$; $D6.3 \geq 3$
Level 4 – Predictive Twin	$D1.1 \geq 4$; $D1.2 \geq 4$; $D1.3 \geq 4$; $D1.4 \geq 4$; $D1.5 \geq 4$	$D2.1 \geq 4$; $D2.2 \geq 4$; $D2.3 \geq 4$	$D3.1 \geq 4$; $D3.2 \geq 4$; $D3.3 \geq 4$; $D3.4 \geq 3$	$D4.1 \geq 3$; $D4.2 \geq 3$; $D4.3 \geq 3$; $D4.4 \geq 2$	$D5.1 \geq 4$; $D5.3 \geq 4$; $D5.4 \geq 4$	$D6.1 \geq 4$; $D6.2 \geq 4$; $D6.3 \geq 4$
Level 5 – Optimising Twin	$D1.1 = 5$; $D1.2 = 5$; $D1.3 = 5$; $D1.4 = 5$; $D1.5 = 5$	$D2.1 = 5$; $D2.2 = 5$; $D2.3 = 5$	$D3.1 = 5$; $D3.2 = 5$; $D3.3 = 5$; $D3.4 \geq 4$	$D4.1 = 5$; $D4.2 \geq 4$; $D4.3 \geq 4$; $D4.4 = 5$	$D5.1 = 5$; $D5.2 \geq 4$; $D5.3 = 5$; $D5.4 = 5$	$D6.1 = 5$; $D6.2 = 5$; $D6.3 = 5$

Part E

Evaluate Artefact

9
Evaluation from Schiphol and YVR

This chapter answers the following sub-question:

SQ4 - To what extent is the digital twin maturity assessment framework applicable and useful for airport turnaround operations?

This chapter evaluates the practical value of the maturity assessment framework. The focus is not on determining Schiphol's maturity again, because the demonstration results were already presented in Chapter 8. Instead, the chapter examines whether the framework was understandable, applicable, and useful for discussing digital twin maturity in airport turnaround operations.

The evaluation is structured in two parts. The first part evaluates the framework based on the Schiphol demonstration and the post-demonstration questionnaire. The second part uses the YVR expert interview to explore whether the framework is also meaningful in another airport context. Together, these evaluations assess both the practical usefulness of the framework in its main case context and its potential applicability beyond Schiphol.

9.1. Schiphol evaluation after demonstration

The Schiphol evaluation focuses on what the demonstration revealed about the framework itself. The scoring results of the Schiphol case were already presented in Chapter 8. Therefore, this section does not repeat the full maturity assessment, but reflects on how the participants used, interpreted, and evaluated the framework during and after the demonstration. The evaluation is organised around four themes that emerged from the transcript and questionnaire responses: the usefulness of the framework for structuring discussion, the clarity of the maturity levels, the distinction between internal maturity and external implementation constraints, and the usefulness of the framework for identifying development steps.

The Schiphol evaluation should therefore be read as an evaluation of the framework's practical value in use. The main question is not whether the final Schiphol score was high or low, but whether the framework helped participants reason about digital twin maturity in a structured and meaningful way.

9.1.1. Structured discussion and awareness

The Schiphol evaluation shows that the framework was useful for structuring a discussion about digital twin maturity. During the demonstration session, participants used the framework to discuss not only the current maturity of Wilbur and related turnaround capabilities, but also the conditions that influence further development. The framework helped make the discussion more explicit by separating data infrastructure, system integration, predictive capability, decision support, operational visibility, and human interaction.

This was especially visible in the discussion about uneven maturity. One participant noted that Wilbur cannot be assessed as one uniform system, because different parts of the system have developed at different speeds:

> *"Some things are not mature at all, while others are already very advanced."* (I7, S4, translated)

This supports the use of separate dimensions and indicators. The framework made it possible to show where maturity is already present and where important capabilities are still missing. This is important because a single overall score would hide these differences.

This was also confirmed in the questionnaire responses. One participant stated that the framework helped the discussion on the maturity of the digital twin, especially by showing where the focus should be to improve the model. Another participant indicated that the most useful part was not one specific dimension, but the fact that the framework created a conversation and a common understanding. This suggests that the framework is useful as a reflective discussion tool, rather than only as a scoring instrument.

The evaluation also showed that the framework can create awareness of uneven maturity. Participants noted that some parts of Wilbur are relatively mature, while other parts are less developed. For example, some functionalities already support advanced insight, while other parts still depend on incomplete data or limited stakeholder integration. This means that the framework is most useful when the assessment scope is clearly defined before scoring. In the demonstration, the scope was therefore narrowed to the airside turnaround process.

9.1.2. Clarity of the maturity levels

The maturity levels were generally understandable, but the demonstration showed that some distinctions need careful explanation. In particular, participants questioned the difference between Level 2, Digital Shadow, and Level 3, Monitoring Twin. The discussion showed that the wording should make the transition from a mostly one-way information display towards a more interactive and shared monitoring environment clearer.

Participants also discussed the distinction between prediction, recommendation, workflow action, and autonomous control. This was especially relevant for Level 5. Some participants noted that if a system starts to initiate decisions, it may move beyond a digital twin and towards an autonomous system. The evaluation therefore shows that the framework should clearly distinguish between a digital twin that recommends or supports action and a system that autonomously controls operations.

The questionnaire responses supported this point. One participant indicated that the maturity levels and scoring descriptions were generally clear, but also noted that there can be overlap between levels and that the progression is not always simply "$1 < 2 < 3$". This suggests that the maturity levels are useful, but should be interpreted as a structured development path rather than as a fully linear scale.

9.1.3. Internal maturity and external implementation constraints

The distinction between internal maturity and external implementation constraints was considered useful, but it required explanation during the session. Some participants initially questioned whether constraints such as data access, commercial interests, decision rights, and workflows were truly external, because they are still part of the airport organisation. The explanation that these constraints are external to the digital twin itself helped clarify the distinction.

This distinction was useful because it showed that digital twin maturity is not only about what the system can technically do. It also depends on whether the operational environment allows these capabilities to be used. For example, participants discussed that safety rules, data ownership, role-based access, and workflow responsibilities can influence the use of higher maturity capabilities, even when the technical foundation is already partly present.

The questionnaire responses also supported this distinction. One participant explained that the distinction between internal maturity and external implementation constraints made sense because it concerns a different level of influence. Another participant identified data and safety as difficult parts of the framework, because Schiphol does not own all relevant data and because Wilbur does not have a safety-critical service level agreement. This confirms that the external constraints are relevant for interpreting the maturity score.

9.1.4. Usefulness for identifying development steps

The framework was considered useful for identifying possible development steps. Participants explained that the model can help show where the digital twin currently stands and which conditions would need to improve before a higher maturity level can be reached. The model therefore supports a structured discussion about development priorities.

The demonstration also showed that the framework can function as a shared reference point. One participant explained that the value of the framework lies in creating a common basis for comparison:

> *"What I like is that the axes on which we assessed this exist, because otherwise: what do you compare it with?"* (I7, S2, translated)

This indicates that the framework helped make maturity discussable by giving participants a shared structure for comparing capabilities and identifying possible improvement areas.

At the same time, participants emphasised that a higher maturity level is not automatically better in every situation. A lower score may reflect a deliberate scope choice, an organisational boundary, or a safety limitation rather than a weakness of the system. For example, limited integration of handler resources does not necessarily mean that Wilbur is poorly designed. It may also reflect the role of Schiphol in the turnaround process and the fact that ground handlers manage parts of the operation themselves.

This point was also made explicitly during the demonstration. One participant noted that the framework should not imply that the entire system must always develop towards the highest level:

> *"It is not necessarily the case that everything has to go to level 5."* (I7, S4, translated)

This reinforces that the framework should be used to connect maturity development to operational needs, rather than to present maturity progression as an end in itself.

The questionnaire responses confirm this interpretation. One participant stated that the framework helped identify areas for improvement and that scoring individual elements helped translate detailed observations into a helicopter view of maturity. However, another participant noted that the next level in the model is not automatically a requirement for Schiphol. According to this response, development should be related to the benefit of reaching a higher level and to concrete operational problems or business requests.

9.1.5. Conclusion of the Schiphol evaluation

Overall, the Schiphol evaluation shows that the framework is useful for reflection, comparison, and prioritisation. It helped participants discuss digital twin maturity in a structured way, identify differences between dimensions, and reflect on the conditions needed for further development. The framework was especially useful because it made clear that maturity depends not only on technical capability, but also on scope, data access, stakeholder roles, workflow integration, and safety boundaries.

The evaluation also shows that the framework should be applied as a structured discussion and decision-support tool rather than as a simple judgement tool. The scope of the assessment should be defined before scoring, and the results should be interpreted together with the underlying dimension scores and operational context. In this way, the framework can help airports determine where digital twin capabilities currently stand, what level is desirable for a specific operational problem, and which development steps are needed to move forward.

9.2. YVR expert evaluation for transferability

The second part of the artefact evaluation consisted of a semi-structured expert interview with a representative from Vancouver International Airport (YVR). The purpose of this interview was not to conduct a full maturity assessment of YVR, but to evaluate whether the framework is understandable, relevant, and potentially transferable to an airport context outside Schiphol. This was particularly relevant because YVR differs from Schiphol in the role of the airport operator in turnaround operations and in the maturity of turnaround-related digital twin capabilities. The YVR transferability evaluation materials are provided in Appendix G.

9.2.1. YVR turnaround context

The interviewee works as an airport flow manager and described the role as monitoring airport processes from curb to cloud. This includes passenger flows, baggage processes, gates, airside flows, traffic, parking, check-in, security, customs, and capacity constraints. In relation to turnaround operations, YVR mainly has a monitoring, coordination, and gate-planning role. The airport schedules gates and monitors whether aircraft are likely to depart on time, but it does not directly execute or control most turnaround activities.

These activities are largely performed by airlines, ground handlers, fuelers, caterers, and other external service providers (I8, Y).

This makes YVR a useful external evaluation context for the framework. In contrast to Schiphol, where the demonstration focused on an existing and relatively more developed turnaround-related digital environment, YVR is still at an earlier stage for turnaround-specific digital twin development. At the same time, YVR already uses digital twin capabilities in other parts of the airport, especially for passenger flow, queue monitoring, terminal capacity, and situational awareness. This contrast made it possible to explore whether the framework could also be understood and applied in a context where digital twin maturity differs across airport domains (I8, Y).

9.2.2. Understanding and relevance of the framework

The interview showed that the framework was understandable to an external airport stakeholder. Unlike the Schiphol participants, the YVR interviewee had not been involved in the earlier development or demonstration of the framework. Nevertheless, the interviewee understood the framework as a way to assess the maturity of digital twin use in turnaround operations, ranging from basic visualisation and manual interpretation towards more advanced predictive and optimisation capabilities:

> *"You're kind of looking at the maturity of where you are in terms of using technology for turnaround."* (I8, Y)

This supports the clarity and communicability of the framework. The model could be understood by an external airport stakeholder without prior involvement in the design process.

The interviewee also considered the framework relevant for YVR. The maturity levels and dimensions were seen as a useful way to structure discussion about where an airport currently stands and what would be needed to develop further. In particular, the interview confirmed the relevance of data availability, system integration, predictive capability, stakeholder collaboration, operational visibility, and decision support as important elements of turnaround-related digital twin maturity.

9.2.3. Relevance of the internal-external distinction

The YVR interview confirmed the relevance of distinguishing between internal digital twin maturity and external implementation constraints. The interviewee explained that YVR is investing in technology and already has digital twin capabilities in the airport environment. However, turnaround-specific maturity is still strongly influenced by external conditions such as data sharing, stakeholder collaboration, commercial sensitivity, decision rights, and operational responsibility.

This was especially clear in relation to turnaround data. YVR depends heavily on airlines, ground handlers, and other partners for information about the status of the turnaround. Some passenger data is already shared by one airline, but this data is not yet available live and does not cover all airlines. The interviewee also explained that information such as passenger loads, connection data, baggage status, fuelling status, catering status, and boarding progress would be important for a more mature turnaround digital twin, but is not yet consistently available to the airport.

The interviewee emphasised that technological development alone is not sufficient. Further maturity depends on whether airlines and ground handlers are willing and able to collaborate and share operational data:

> *"Once you get that collaboration and agreement, then the maturity can ramp up quite quickly from that point."* (I8, Y)

This supports the framework's assumption that technical capability alone is not sufficient to determine practical maturity. Even when an airport has an advanced digital foundation, the use of a turnaround digital twin depends on whether the relevant data can be accessed, whether stakeholders are willing to share information, and whether the airport has the authority to use digital twin outputs in operational decision-making.

9.2.4. Evaluation of maturity levels and dimensions

The maturity levels were considered logical by the interviewee. For turnaround-specific digital twin maturity, YVR assuned to be positioned between the Digital Shadow and Monitoring Twin levels. Some real-time operational information is available, but YVR does not yet have a complete shared monitoring view of the full turnaround process. For example, the airport can monitor certain operational indicators and use CCTV or other digital systems, but it does not yet have full visibility of activities such as fuelling, baggage loading, catering, boarding progress, or ground handler resource status (I8, Y).

The interviewee also agreed that the Optimising Twin represents the highest maturity level in the framework. At this level, accurate data, prediction, automation, and operational recommendations are combined to support more efficient airport decision-making. The six internal maturity dimensions were also regarded as relevant. The data dimension was considered especially important, because YVR is currently working on data governance, data consistency, and improved internal data structures. The interviewee explained that different departments may currently use or interpret data differently, which shows why a consistent data foundation is necessary before more advanced digital twin capabilities can be developed (I8, Y).

9.2.5. Transferability of the framework

Overall, the YVR evaluation suggests that the framework is transferable beyond the Schiphol case, but not as a fixed benchmark for comparing airports directly. The core structure of the framework, consisting of maturity levels, internal dimensions, and external implementation constraints, was considered applicable to YVR. However, the interpretation of specific scores remains context-dependent, because airports differ in their governance structures, data-sharing arrangements, operational responsibilities, stakeholder relationships, and regulatory environments.

The interviewee summarised this contextual variation as follows:

> *"When you've seen one airport, you've only seen one airport."* (I8, Y)

This statement does not imply that the framework is only applicable to Schiphol. Rather, it indicates that transferability should be understood at the level of the framework's structure, while the interpretation of assessment outcomes requires attention to the local airport context. For example, a low score on data completeness or operational decision optimisation may reflect limited airport authority over ground handlers, rather than a lack of technical maturity in the digital twin itself.

This interpretation is supported by the interviewee's assessment of the framework:

> *"It is transferable with adaptations, because every airport operates in a different way."* (I8, Y)

In the YVR case, this is particularly relevant because the airport mainly has a monitoring and gate-planning role, while many turnaround activities are executed by airlines, ground handlers, fuelers, caterers, and other external parties. Maturity limitations may therefore arise not only from the digital twin's internal capabilities, but also from the airport's ability to access relevant data, influence operational decisions, and integrate external stakeholders into shared workflows.

The YVR evaluation therefore adds an important nuance to the framework's transferability. The maturity levels, internal dimensions, and external constraints appear relevant beyond Schiphol. However, the scores should always be interpreted in relation to the local airport context. A score does not only reflect the technical maturity of the digital twin, but also the airport's role in the turnaround process, its access to data, its relationships with stakeholders, and the safety requirements that apply. The framework should therefore be used to support context-sensitive discussion, rather than to rank airports directly.

9.2.6. Conclusion of the YVR evaluation

The YVR expert evaluation supports the broader applicability of the framework. The framework was understood without prior involvement, the maturity levels were considered logical, and the internal-external distinction was confirmed as relevant. The interview also showed that the framework can be useful in an airport context where digital twin capabilities are more advanced in some domains, such as passenger flow and terminal monitoring, than in turnaround operations. At the same time, the evaluation shows that the framework should be applied as a context-sensitive assessment instrument rather than as a fixed universal model (I8, Y).

9.3. Chapter conclusion

This chapter evaluated whether the maturity assessment framework is understandable, applicable, and useful for airport turnaround operations. The Schiphol evaluation showed that the framework is useful as a structured discussion and reflection tool. It helped participants discuss maturity across separate dimensions, identify uneven maturity, and connect current scores to possible development steps. It also showed that the framework should be applied to a clearly defined operational scope and that scores should be interpreted in relation to organisational boundaries, data access, decision rights, and safety constraints.

The YVR evaluation showed that the framework can also be understood and discussed outside the Schiphol context. The interview confirmed that the distinction between internal maturity and external implementation constraints is relevant in another airport setting. However, the YVR evaluation also confirmed that transferability depends on context-specific interpretation. Airports differ in their operational responsibilities, stakeholder relationships, and data-sharing arrangements.

Overall, the evaluation indicates that the framework is applicable and useful as an exploratory maturity assessment instrument for airport turnaround digital twins. It provides a structured basis for discussing current maturity, identifying limiting conditions, and guiding future digital twin development in airport turnaround operations.

Part F

Discussion and Conclusion

10 Discussion and conclusion

This chapter discusses the findings of this thesis and answers the main research question. It first presents the answer to the research question by summarising the final maturity assessment framework. It then interprets the main findings, and discusses the theoretical and practical contributions. Finally, the chapter reflects on the main limitations of the research and presents directions for future research.

10.1. Answer to the main research question

The main research question of this thesis was:

Which maturity assessment framework can guide the development of digital twins in airport turnaround operations?

This thesis answers the research question by proposing a context-specific digital twin maturity assessment framework for airport turnaround operations. The framework is designed to evaluate the current maturity of an airport turnaround digital twin and to guide further development by identifying capability gaps and possible next steps. It combines maturity levels, internal maturity dimensions, indicators, external implementation constraints, and structured scoring logic.

The framework was developed through four research steps. First, the study identified the functional, structural, and environmental requirements that a maturity assessment framework for airport turnaround digital twins should fulfil. Second, these requirements were translated into maturity levels, internal maturity dimensions, indicators, external implementation constraints, and scoring logic. Third, the framework was demonstrated in the Schiphol context to examine whether it could structure a maturity assessment in a real airport turnaround setting. Fourth, the evaluation phase with Schiphol and Vancouver International Airport (YVR) was used to assess the framework's clarity, usefulness, and potential transferability. Together, these steps led to the final maturity assessment framework presented in this thesis.

The final framework consists of a baseline level below the maturity threshold and five maturity levels: Digital Model, Digital Shadow, Monitoring Twin, Predictive Twin, and Optimizing Twin. These levels describe the development of a digital twin from a static or non-real-time representation of turnaround operations towards a system that can support shared monitoring, prediction, decision support, resource adjustment, task and schedule adjustment, and workflow-ready operational action. In contrast to some existing maturity models, the highest level is not defined as a fully autonomous twin. Instead, Level 5 is defined as an Optimizing Twin, because the framework focuses on the digital twin's ability to translate predictive insights into structured operational recommendations and actions. Automation can be part of this level, for example when actions are proposed, routed, or initiated through connected systems, but it is not treated as a separate final maturity stage. This keeps the framework aligned with the airport turnaround context, where operational decisions and physical execution remain embedded in existing human, organisational, workflow, accountability, and safety boundaries.

To assess maturity in more detail, the framework includes six internal maturity dimensions: Data Infrastructure, System Integration, Predictive and Analytical Capability, Operational Decision Optimisation, Operational Situation Visibility, and Human Interaction and Trust. These dimensions are operationalised through indicators that translate each dimension into assessable digital twin capabilities. Together, the dimensions and indicators make digital twin maturity measurable as a multidimensional construct rather than as one single overall score. This is important because digital twin capabilities may develop unevenly across dimensions. For example, an airport may already have strong operational visibility and data infrastructure, while predictive analytics or workflow-ready decision support are still less mature.

The framework can produce an overall internal maturity level, but this level is not based on a simple average. Instead, the indicator scores are combined through threshold logic. The dimensions and indicators were developed and refined through literature and expert feedback, while the detailed 0–5 scoring descriptions and threshold values should be interpreted more cautiously. These scoring descriptions support structured application and make the assessment process more transparent, but they are best understood as a structured application proposal rather than as a fully validated quantitative measurement instrument.

A central design choice of the framework is the distinction between internal maturity and external implementation constraints. Internal maturity assesses what the digital twin is capable of delivering, while external constraints assess whether those capabilities can realistically be used in the airport context. This distinction is important because technical capability does not automatically lead to operational use. A digital twin may be able to support prediction or optimisation, but its use can still be limited by data access, commercial sensitivity, decision rights, workflow integration, accountability structures, or safety boundaries.

10.2. Interpretation of the findings

The findings show that the framework provides a useful starting point for discussing digital twin maturity in airport turnaround operations. Rather than treating digital twin maturity as a general or abstract concept, the framework makes it possible to identify where digital capabilities currently lie within the turnaround context and which development steps may be needed next. This is especially relevant because airport turnaround operations involve interdependent actors, data sources, systems, responsibilities, and operational decisions.

A key finding is that maturity in this context should not be understood as one single overall level or as a strictly linear development path. The Schiphol demonstration showed that digital twin capabilities can develop unevenly across different dimensions. For example, an airport may have relatively strong capabilities in data infrastructure, system integration, and operational visibility, while predictive analytics, decision support, or workflow-ready optimisation are still less developed. At the same time, some decision-support capabilities may already exist before predictive capabilities are fully mature, for example when operational systems support recommendations based on rules, current-state information, or expert logic rather than advanced prediction. The maturity levels should therefore be interpreted as a structured development path rather than a rigid sequence.

The results also show that the distinction between internal maturity and external implementation constraints is necessary for interpreting maturity in practice. Internal maturity describes what the digital twin can do. External constraints describe whether these capabilities can realistically be used in the airport context. This means that maturity is not only a technical property of the digital twin itself. It is also shaped by the airport's role in the turnaround process, the willingness and ability of stakeholders to share data, commercial sensitivity, decision rights, workflow arrangements, and safety requirements.

The Schiphol demonstration and evaluation showed that the framework is especially useful as a structured discussion and diagnostic tool. It helped participants move beyond the question of whether Schiphol has a digital twin and instead discuss which capabilities are mature, which are less mature, and which internal or external conditions limit further development. The framework therefore supports both assessment and learning. It does not only assign a maturity level, but also helps identify strengths, bottlenecks, and development priorities.

The YVR evaluation adds an important nuance to the transferability of the framework. The framework was understandable and relevant in a different airport context, and the maturity levels, internal dimensions, and distinction between internal maturity and external constraints were considered useful. However, the interpretation of specific scores remains context-dependent. Airports differ in their governance structures, operational responsibilities, stakeholder relationships, data-sharing arrangements, and regulatory conditions. The framework should therefore be transferred as an assessment structure, not as a fixed scoring template for direct comparison between airports without contextual interpretation.

10.3. Academic contribution

This thesis contributes to digital twin maturity literature by adapting existing maturity-model logic to the specific context of airport turnaround operations. Existing digital twin maturity models provide useful concepts for structuring maturity across levels, dimensions, indicators, and assessment logic. For example, Uhlenkamp et al. (2022) assess digital twins across categories such as data, integration, control, model, computing capabilities, and human-machine interface, while L. Chen et al. (2021) emphasise that maturity should include purpose, trust, and function rather than only technical capability. However, these models are mostly developed for domains such as production, logistics, asset management, infrastructure, or buildings (L. Chen et al., 2021; Z.-S. Chen et al., 2024; Li et al., 2024; Uhlenkamp et al., 2022). They therefore provide a useful foundation, but do not directly address the multi-actor and operationally distributed nature of airport turnaround operations.

The first academic contribution is the development of a domain-specific maturity assessment framework for airport turnaround digital twins. Turnaround operations are time-sensitive, safety-sensitive, and dependent on coordination between multiple independent stakeholders, including the airport operator, airlines, ground handlers, and air traffic-related actors. A maturity framework for this context therefore needs to assess more than technical digital twin development alone. It also needs to consider whether the digital twin supports shared operational visibility, data integration, prediction, decision support, stakeholder coordination, and human interaction. In this way, the research translates general digital twin maturity thinking into an assessment framework for a specific airport operations context.

This contribution also complements existing airport digital twin studies. Studies such as Conde et al. (2022), Luo et al. (2024), Lu et al. (2025), and Attar et al. (2024) show how digital twins can support airport turnaround visualisation, vehicle scheduling, automated apron operations, and airside optimisation. These studies demonstrate the potential value of digital twins in airport operations, but they mainly focus on digital twin applications, architectures, simulations, or optimisation methods. This thesis takes a different perspective by asking how airports can assess whether such capabilities are mature enough to support turnaround coordination in practice. The contribution therefore lies in providing an assessment and development framework rather than another technical implementation.

The second academic contribution is the distinction between internal digital twin maturity and external implementation constraints. Existing maturity models often focus on the technical or functional capabilities of the digital twin itself. However, in airport turnaround operations, practical maturity also depends on data availability, stakeholder collaboration, commercial sensitivity, decision rights, workflow integration, accountability structures, and safety requirements. By separating what the digital twin is capable of delivering from what the airport context allows it to do in practice, the framework extends digital twin maturity thinking towards a more socio-technical understanding of maturity. This distinction helps explain why technically advanced capabilities may still have limited operational use when external implementation conditions restrict their application (Haraguchi et al., 2024; Moon et al., 2025).

The third academic contribution concerns the interpretation of advanced maturity. Several digital twin maturity models describe progression towards prescriptive or autonomous twins (Li et al., 2024; Y. Liu et al., 2024; Manickam et al., 2023). In this framework, full autonomy is not treated as a separate final maturity level. Instead, the highest level is defined as an Optimizing Twin, where predictive insights are translated into optimisation-oriented decision support, such as operational recommendations, resource adjustments, task and schedule adjustments, or workflow-ready actions. Automation can be part of this level, but it is not treated as the main endpoint of maturity. This interpretation fits the airport turnaround context, where operational decisions and physical execution remain embedded in human, organisational, workflow, accountability, and safety boundaries. It also aligns with literature that emphasises user experience, trust, explainability, and human interaction in digital twin use (K. Chen et al., 2024; Manickam et al., 2023).

10.4. Practical contribution

The practical contribution of this thesis is that it provides airport operators with a structured way to assess whether digital twin capabilities are mature enough to create operational value in airport turnaround operations. This is important because digital twin technology only supports airport performance when it moves beyond a digital representation or dashboard and can support reliable data use, integrated operational visibility, prediction, and decision support. ACI EUROPE argues that airports require a clear overview of how the schedule is being met, with particular attention to the aircraft turnaround process, and that airports should strengthen real-time monitoring of aircraft turnaround, queues, and delays through IT, AI, and digital tools (ACI EUROPE, 2025). Similarly, EUROCONTROL describes Airport Collaborative Decision Making as a concept that improves airport efficiency and resilience through transparent collaboration and the exchange of accurate and timely information between airport operators, aircraft operators, ground handlers, ATC, and the Network Manager, with a specific focus on turn-round and pre-departure processes (EUROCONTROL, n.d.). The framework developed in this thesis supports these practical needs by helping airports assess whether their digital twin capabilities are mature enough to support operational visibility, coordination, prediction, and decision support in turnaround operations.

Without such a maturity assessment, digital twin development risks becoming technology-driven rather than problem-driven. Airports may invest in visible digital tools, data platforms, dashboards, or predictive models without knowing whether the underlying capabilities needed for operational use are actually present. Important weaknesses may then remain hidden, such as incomplete data, fragmented systems, limited predictive capability, unclear decision rights, weak workflow integration, limited user trust, or restricted data sharing between stakeholders. In that case, the digital twin may appear advanced, while still having limited value for real turnaround coordination.

The framework addresses this problem by moving the assessment beyond the question of whether an airport “has” a digital twin. Instead, it helps airport operators identify what type of digital twin capability exists, how mature this capability is across different dimensions, and which internal or external conditions limit further development. The framework therefore produces more than one overall maturity label. It creates a maturity profile that shows strengths and weaknesses across data infrastructure, system integration, predictive and analytical capability, operational decision optimisation, operational situation visibility, and human interaction and trust. This makes the framework useful as a diagnostic tool for understanding the current state of digital twin development in airport turnaround operations.

A second practical value is that the framework translates maturity assessment into development priorities. Airport digital twin capabilities may develop unevenly. For example, an airport may already have strong capabilities in data infrastructure, system integration, or operational visibility, while predictive analytics, scenario analysis, or operational decision support are still less mature. By making these differences explicit, the framework helps airport operators identify where further development is most needed. It can therefore support more targeted discussions about investment priorities, digital development ambitions, data quality improvements, system integration, predictive analytics, and decision-support functionality.

A third practical value is that the framework makes implementation constraints explicit. The distinction between internal maturity and external implementation constraints helps airport operators understand whether further development is mainly limited by the digital twin itself or by the surrounding implementation context. For example, a digital twin may be technically capable of supporting prediction or optimisation, while its operational use is still limited by data access, commercial sensitivity, decision rights, workflow integration, accountability structures, or safety requirements. This makes the framework useful not only for technical development planning, but also for structuring discussions about governance, stakeholder collaboration, and implementation feasibility.

The Schiphol demonstration and YVR evaluation support this practical contribution, but also define its boundaries. The Schiphol demonstration showed that the framework is especially useful as a structured discussion and reflection tool, because it helped participants discuss digital twin maturity across data, integration, prediction, decision support, operational visibility, and human interaction. The YVR evaluation indicated that the framework was understandable and relevant in another airport context. At the same time, the evaluation showed that the framework should not yet be treated as a universally validated quantitative measurement model or as a fixed benchmarking instrument for comparing airports. Its current practical value lies in supporting structured assessment, stakeholder discussion, and development planning, while broader validation is needed before the framework can be used for direct comparison across airports.

10.5. Limitations and future research

This study has several limitations that should be considered when interpreting the results. These limitations do not undermine the value of the framework, but they indicate how the framework should be used and where future research is needed.

First, the framework was developed and demonstrated mainly in the Schiphol context. Schiphol is a relevant and information-rich case because it already uses A-CDM and has developed digital technologies such as Wilbur and Deep Turnaround. However, this also means that the framework was tested in a relatively mature airport environment. Other airports may have different organisational structures, data-sharing arrangements, operational priorities, or levels of digital development. YVR was included as an external evaluation case, but not as a full second case study. The YVR evaluation showed that the framework was understandable and relevant in another airport context, but the framework was not fully applied and scored at YVR in the same way as at Schiphol. Future research should therefore apply the full framework, including the scoring process, at multiple airports with different levels of digital development. This would make it possible to assess whether the maturity levels, dimensions, indicators, and external constraints remain useful across airports that differ in size, governance structure, operational complexity, and digital twin maturity.

Second, the scoring logic requires further validation. The maturity levels, dimensions, and indicators were developed and refined through literature and expert feedback, but the detailed 0–5 scoring descriptions were mainly formulated by the researcher as an operationalisation of these indicators. This means that the indicators themselves were reviewed, while the exact score descriptions, threshold values, and scoring chart were not validated with the same depth. In addition, maturity assessment still depends partly on expert judgement. Different experts may interpret certain maturity levels or indicators differently, especially when capabilities develop unevenly. Future research should therefore test whether different experts interpret the 0–5 scores consistently, whether the score descriptions are clear enough in practice, and whether the thresholds lead to credible maturity outcomes. Comparing assessments made by different experts, examining inter-rater reliability, and exploring whether weighting of dimensions or indicators is needed would strengthen the framework as a more quantitative assessment instrument.

third, digital twin maturity was assessed at one point in time. The current study provides a framework for assessing the maturity of an airport turnaround digital twin at a specific moment, but it does not yet show how maturity develops over time. Future research could apply the framework longitudinally during the development of an airport digital twin. This would show whether maturity actually develops in the sequence suggested by the framework, or whether some dimensions develop faster than others. Such research would also help explain how airports move from monitoring towards prediction, decision support, and optimisation in practice.

Finally, the literature review may not have captured all relevant work on airport digital twins. Digital twin is a broad and sometimes inconsistently used term. Some systems may function like digital twins without being explicitly labelled as such. Wilbur is an example of this: it was not necessarily developed or described as a digital twin from the beginning, but it has several characteristics that are relevant for digital twin maturity. This means that some relevant studies on operational platforms, real-time airport systems, simulation environments, decision-support systems, or collaborative airport technologies may not have appeared in the literature search if they did not use the term digital twin. Future research could therefore broaden the literature base by including studies that describe digital-twin-like systems without explicitly using the term digital twin. This could further strengthen the theoretical foundation of airport digital twin maturity assessment.

Overall, these limitations mean that the framework should be seen as an initial maturity assessment framework rather than a universally validated model. It provides a structured basis for assessing and discussing digital twin maturity in airport turnaround operations, but broader validation is needed across more airports, stakeholder groups, and operational contexts.

10.6. Final remarks

This research started with the aim of developing a framework that could assess the maturity of airport turnaround digital twins. During the research, it became clear that this maturity is difficult to capture in one single score. Airport turnaround operations are complex, time-sensitive, and distributed across multiple actors, systems, responsibilities, and safety boundaries. A maturity assessment therefore needs to show differences between dimensions and explain the conditions that limit or enable further development.

The final framework addresses this complexity by combining maturity levels with internal dimensions, indicators, external implementation constraints, and scoring logic. Its main value is not only that it can assign a maturity level, but also that it can structure a meaningful discussion about digital twin development. In this way, the framework provides a starting point for airports that want to understand their current digital twin capabilities and decide which development steps are relevant, realistic, and valuable in their own operational context.

In conclusion, the maturity assessment framework designed in this thesis evaluates and guides the development of airport turnaround digital twins by connecting technical digital twin capabilities with the operational and organisational conditions under which these capabilities can be used. It provides airport operators with a structured way to assess where their digital twin capabilities currently stand, where maturity differs across dimensions, and which development steps are needed to move towards more integrated, predictive, and decision-supporting digital twin use in turnaround operations.

Bibliography

Abdelalim, A. M., Essawy, A., Sherif, A., Salem, M., Al-Adwani, M., & Abdullah, M. S. (2025). Optimizing Facilities Management Through Artificial Intelligence and Digital Twin Technology in Mega-Facilities. *Sustainability*, *17*(5), 1826. https://doi.org/10.3390/su17051826

ACI EUROPE. (2025). *Airport Performance Position Paper 2025* (Position paper). ACI EUROPE. https://www.aci-europe.org/downloads/resources/ACI%20EUROPE%20Performance%20Position%20Paper%202025.pdf

ACI World. (2026, January). ACI World releases global airport traffic forecasts as long-term demand growth continues to reshape aviation. https://aci.aero/2026/01/28/aci-world-releases-global-airport-traffic-forecasts-as-long-term-demand-growth-continues-to-reshape-aviation/

Adams, W. C. (2015, August). Conducting Semi-Structured Interviews. In K. E. Newcomer, H. P. Hatry, & J. S. Wholey (Eds.), *Handbook of Practical Program Evaluation* (1st ed., pp. 492–505). Wiley. https://doi.org/10.1002/9781119171386.ch19

Archibald, M. M., Ambagtsheer, R. C., Casey, M. G., & Lawless, M. (2019). Using Zoom Videoconferencing for Qualitative Data Collection: Perceptions and Experiences of Researchers and Participants. *International Journal of Qualitative Methods*, *18*, 1609406919874596. https://doi.org/10.1177/1609406919874596

Attar, A., Babaee, M., Raissi, S., & Nojavan, M. (2024). Airside Optimization Framework Covering Multiple Operations in Civil Airport Systems with a Variety of Aircraft: A Simulation-Based Digital Twin. *Systems*, *12*(10), 394. https://doi.org/10.3390/systems12100394

Becker, J., Knackstedt, R., & Pöppelbuß, J. (2009). Developing Maturity Models for IT Management: A Procedure Model and its Application. *Business & Information Systems Engineering*, *1*(3), 213–222. https://doi.org/10.1007/s12599-009-0044-5

Cappellazzo, V. (2025). What is A-CDM. https://www.icao.int/sites/default/files/APAC/Meetings/2025/2025%20Capacity%20Assessment%20WS/Day%201/SP-04-What-is-A-CDM-by-EUROCONTROL.pdf

Chen, K., Nadirsha, T. N., Lilith, N., Alam, S., & Svensson, Å. (2024). Tangible digital twin with shared visualization for collaborative air traffic management operations. *Transportation Research Part C: Emerging Technologies*, *161*, 104546. https://doi.org/10.1016/j.trc.2024.104546

Chen, L., Xie, X., Lu, Q., Parlikad, A. K., Pitt, M., & Yang, J. (2021). Gemini Principles-Based Digital Twin Maturity Model for Asset Management. *Sustainability*, *13*(15), 8224. https://doi.org/10.3390/su13158224

Chen, Z.-S., Chen, K.-D., Xu, Y.-Q., Pedrycz, W., & Skibniewski, M. J. (2024). Multiobjective optimization-based decision support for building digital twin maturity measurement. *Advanced Engineering Informatics*, *59*, 102245. https://doi.org/10.1016/j.aei.2023.102245

Conde, J., Munoz-Arcentales, A., Romero, M., Rojo, J., Salvachúa, J., Huecas, G., & Alonso, Á. (2022). Applying digital twins for the management of information in turnaround event operations in commercial airports. *Advanced Engineering Informatics*, *54*, 101723. https://doi.org/10.1016/j.aei.2022.101723

Digital Twin Consortium. (2024, October). *Digital Twin Value Overview and Use Cases for Airport Operations* (tech. rep.) (White paper). Digital Twin Consortium.

Durmaz, R., & Göçmen, F. (2025). Usage Of Dıgıtal Twın Technologıes In Aırports: The Case Of Istanbul Aırport. *EAI Endorsed Transactions on Tourism, Technology and Intelligence*, *2*(4). https://doi.org/10.4108/eettti.9626

EUROCONTROL. (n.d.). Airport collaborative decision-making (A-CDM). Retrieved March 9, 2026, from https://www.eurocontrol.int/concept/airport-collaborative-decision-making

Evler, J., Lindner, M., Fricke, H., & Schultz, M. (2022). Integration of turnaround and aircraft recovery to mitigate delay propagation in airline networks. *Computers & Operations Research*, *138*, 105602. https://doi.org/10.1016/j.cor.2021.105602

Fereday, J., & Muir-Cochrane, E. (2006). Demonstrating Rigor Using Thematic Analysis: A Hybrid Approach of Inductive and Deductive Coding and Theme Development. *International Journal of Qualitative Methods*, *5*(1), 80–92. https://doi.org/10.1177/160940690600500107

Flyvbjerg, B. (2006). Five Misunderstandings About Case-Study Research. *Qualitative Inquiry*, *12*(2), 219–245. https://doi.org/10.1177/1077800405284363

Fuchte, J., Nagel, B., & Gollnick, V. (2012). Twin Aisle for Short Range Operations - An Economically Attractive Alternative? *12th AIAA Aviation Technology, Integration, and Operations (ATIO) Conference and 14th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference*. https://doi.org/10.2514/6.2012-5408

Gourisetti, S. N. G., Sebastian Cardenas, D., Akshay, R., Reid, W., Michael, M., Rene, K., & Cali, U. (2025). Cybersecurity, digital privacy, and modeling aspects of digital twins. In *Digital Twin Technology for the Energy Sector* (pp. 47–93). Elsevier. https://doi.org/10.1016/B978-0-443-14070-9.00003-2

Guest, G., Bunce, A., & Johnson, L. (2006). How Many Interviews Are Enough?: An Experiment with Data Saturation and Variability. *Field Methods*, *18*(1), 59–82. https://doi.org/10.1177/1525822X05279903

Halpern, N., Budd, T., Suau-Sanchez, P., Bråthen, S., & Mwesiumo, D. (2021). Conceptualising airport digital maturity and dimensions of technological and organisational transformation. *Journal of Airport Management*, *15*(2), 182. https://doi.org/10.69554/MZJB2664

Haraguchi, M., Funahashi, T., & Biljecki, F. (2024). Assessing governance implications of city digital twin technology: A maturity model approach. *Technological Forecasting and Social Change*, *204*, 123409. https://doi.org/10.1016/j.techfore.2024.123409

Hartmann, M., & Wimmer, M. A. (2026). Maturity Models for Digital Twins in Smart Cities – Literature Review and Comparison [Series Title: Lecture Notes in Computer Science]. In I. Lindgren, M. P. Rodríguez Bolívar, M. Janssen, E. Loukis, F. Mureddu, P. Panagiotopoulos, G. Viale Pereira, & E. Tambouris (Eds.), *Electronic Government* (pp. 267–283, Vol. 15944). Springer Nature Switzerland. https://doi.org/10.1007/978-3-032-01589-1_17

Hennink, M. M., Kaiser, B. N., & Marconi, V. C. (2017). Code Saturation Versus Meaning Saturation: How Many Interviews Are Enough? *Qualitative Health Research*, *27*(4), 591–608. https://doi.org/10.1177/1049732316665344

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design Science in Information Systems Research.

Iliuță, M.-E., Moisescu, M.-A., Pop, E., Ionita, A.-D., Caramihai, S.-I., & Mitulescu, T.-C. (2024). Digital Twin—A Review of the Evolution from Concept to Technology and Its Analytical Perspectives on Applications in Various Fields. *Applied Sciences*, *14*(13), 5454. https://doi.org/10.3390/app14135454

Johannesson, P., & Perjons, E. (2014). *An Introduction to Design Science*. Springer International Publishing. https://doi.org/10.1007/978-3-319-10632-8

Kitzinger, J. (1995). Qualitative Research: Introducing focus groups. *BMJ*, *311*(7000), 299–302. https://doi.org/10.1136/bmj.311.7000.299

Kritzinger, W., Karner, M., Traar, G., Henjes, J., & Sihn, W. (2018). Digital Twin in manufacturing: A categorical literature review and classification. *IFAC-PapersOnLine*, *51*(11), 1016–1022. https://doi.org/10.1016/j.ifacol.2018.08.474

Li, T., Rui, Y., Zhao, S., Zhang, Y., Zhu, H., & Li, X. (2024). A quantitative digital twin maturity model for underground infrastructure based on D-ANP. *Tunnelling and Underground Space Technology*, *146*, 105612. https://doi.org/10.1016/j.tust.2024.105612

Liu, C., Zhang, Y., Chen, Y., Liu, S., Hu, S., Luo, Q., & Chen, L. (2024). Digital Twin-Enabled Delay Diagnosis Traceability and Propagation Process for Airport Flight Ground Service (Y.-A. Tan, Ed.). *International Journal of Intelligent Systems*, *2024*(1), 7458758. https://doi.org/10.1155/int/7458758

Liu, Y., Feng, J., Lu, J., & Zhou, S. (2024). A review of digital twin capabilities, technologies, and applications based on the maturity model. *Advanced Engineering Informatics*, *62*, 102592. https://doi.org/10.1016/j.aei.2024.102592

Lu, J., Pang, T., Lu, X., Ji, J., & Jiang, C. (2025). Harnessing digital twin technology for enhanced aircraft turnaround efficiency. *Transport Economics and Management*, *3*, 334–345. https://doi.org/10.1016/j.team.2025.09.001

Luo, Q., Liu, H., Liu, C., & Deng, Q. (2024). Multi-strategy cooperative scheduling for airport specialized vehicles based on digital twins. *Scientific Reports*, *14*(1), 15533. https://doi.org/10.1038/s41598-024-66350-0

Madni, A. M., Madni, C. C., & Lucero, S. D. (2019). Leveraging Digital Twin Technology in Model-Based Systems Engineering. *Systems*, *7*(1), 7. https://doi.org/10.3390/systems7010007

Manickam, R., Vollmar, J., & Prabhakar, G. M. (2023). User Experience–Digital Twin Maturity Model (UX-DTMM) [Series Title: Smart Innovation, Systems and Technologies]. In A. Chakrabarti & V. Singh (Eds.), *Design in the Era of Industry 4.0, Volume 2* (pp. 877–889, Vol. 342). Springer Nature Singapore. https://doi.org/10.1007/978-981-99-0264-4_72

McMillan, S. S., King, M., & Tully, M. P. (2016). How to use the nominal group and Delphi techniques. *International Journal of Clinical Pharmacy*. https://doi.org/10.1007/s11096-016-0257-x

Mizrak, F., Polat, L., & Tasar, S. A. (2024). Applying Entropy Weighting and 2-Tuple Linguistic T-Spherical Fuzzy MCDM: A Case Study of Developing a Strategic Sustainability Plan for Istanbul Airport. *Sustainability*, *16*(24), 11104. https://doi.org/10.3390/su162411104

Mo, F., Monetti, F. M., Torayev, A., Rehman, H. U., Mulet Alberola, J. A., Rea Minango, N., Nguyen, H. N., Maffei, A., & Chaplin, J. C. (2023). A maturity model for the autonomy of manufacturing systems. *The International Journal of Advanced Manufacturing Technology*, *126*(1-2), 405–428. https://doi.org/10.1007/s00170-023-10910-7

Moon, S., Kim, G., Seo, H., Jun, J., & Park, E. (2025). A Study on Information Strategy Planning (ISP) for Applying Smart Technologies to Airport Facilities in South Korea. *Aerospace*, *12*(7), 595. https://doi.org/10.3390/aerospace12070595

Morgan, D. L. (1996). Focus Groups. *Annual Review of Sociology*, *2*, 129–152. http://www.jstor.org/stable/2083427

Okwir, S., Ulfvengren, P., Angelis, J., Ruiz, F., & Núñez Guerrero, Y. M. (2017). Managing turnaround performance through Collaborative Decision Making. *Journal of Air Transport Management*, *58*, 183–196. https://doi.org/10.1016/j.jairtraman.2016.10.008

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *Journal of Clinical Epidemiology*, *134*, 178–189. https://doi.org/10.1016/j.jclinepi.2021.03.001

Patton, M. Q. (2015). *Qualitative research & evaluation methods: Integrating theory and practice* (Fourth edition). SAGE.

Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A Design Science Research Methodology for Information Systems Research. *Journal of Management Information Systems*, *24*(3), 45–77. https://doi.org/10.2753/MIS0742-1222240302

Qiu, Y. (2024). Digital Twin Based Resource Optimization and Dynamic Decision Making in Air Terminal Yards. *Applied Mathematics and Nonlinear Sciences*, *9*(1), 20240871. https://doi.org/10.2478/amns-2024-0871

Royal Schiphol Group. (2020, November). Wilbur: Data on passenger flows and 'hot pax'. https://www.schiphol.nl/en/innovation/blog/wilbur-data-on-passenger-flows-and-hot-pax

Royal Schiphol Group. (2024, July). Deep Turnaround Become a smooth operator. https://www.schiphol.nl/en/aviation-solutions/deep-turnaround/

Royal Schiphol Group. (2025, November). Vancouver airport implements Deep Turnaround (Press release). Retrieved March 9, 2026, from https://www.schiphol.nl/en/aviation-solutions/blog/vancouver-airport-implements-deep-turnaround-press-release

Saifutdinov, F., Jackson, I., & Tolujevs, J. (2022). An emulation oriented method and tool for test of ground traffic control systems at airports. *Aviation*, *26*(2), 104–111. https://doi.org/10.3846/aviation.2022.16869

Scheibmeir, J., & Malaiya, Y. (2022). Social Media Perspectives on Digital Twins and the Digital Twins Maturity Model. In Z. Lv & E. Fersman (Eds.), *Digital Twins: Basics and Applications* (pp. 73–99). Springer International Publishing. https://doi.org/10.1007/978-3-031-11401-4_8

Schiphol. (2020, November). Turnaround Insights. https://www.schiphol.nl/nl/innovatie/blog/een-voorspelbaar-omdraaiproces-met-turnaround-insights/

Schmidt, M. (2017). A review of aircraft turnaround operations and simulations. *Progress in Aerospace Sciences*, *92*, 25–38. https://doi.org/10.1016/j.paerosci.2017.05.002

Seawright, J., & Gerring, J. (2008). Case Selection Techniques in Case Study Research: A Menu of Qualitative and Quantitative Options. *Political Research Quarterly*, *61*(2), 294–308. https://doi.org/10.1177/1065912907313077

Singh, M., Fuenmayor, E., Hinchy, E., Qiao, Y., Murray, N., & Devine, D. (2021). Digital Twin: Origin to Future. *Applied System Innovation*, *4*(2), 36. https://doi.org/10.3390/asi4020036

Temple, B., & Young, A. (2004). Qualitative Research and Translation Dilemmas. *Qualitative Research*, *4*(2), 161–178. https://doi.org/10.1177/1468794104044430

Uhlenkamp, J.-F., Hauge, J. B., Broda, E., Lutjen, M., Freitag, M., & Thoben, K.-D. (2022). Digital Twins: A Maturity Model for Their Classification and Evaluation. *IEEE Access*, *10*, 69605–69635. https://doi.org/10.1109/ACCESS.2022.3186353

Unity. (n.d.). Vancouver International Airport Digital Twin | Case Study. Retrieved September 24, 2025, from https://unity.com/case-study/vancouver-airport-authority

Vail, S., Churchill, A., Karlsson, J., McInerney, T., Domitrovich, J., Phillips, T., Airport Cooperative Research Program, Transportation Research Board, & National Academies of Sciences, Engineering, and Medicine. (2015, June). *Guidebook for Advancing Collaborative Decision Making (CDM) at Airports* [Pages: 22121]. Transportation Research Board. https://doi.org/10.17226/22121

Van Der Horn, E., & Mahadevan, S. (2021). Digital Twin: Generalization, characterization and implementation. *Decision Support Systems*, *145*, 113524. https://doi.org/10.1016/j.dss.2021.113524

Van Nes, F., Abma, T., Jonsson, H., & Deeg, D. (2010). Language differences in qualitative research: Is meaning lost in translation? *European Journal of Ageing*, *7*(4), 313–316. https://doi.org/10.1007/s10433-010-0168-y

Wu, C.-L., & Caves, R. E. (2000). Aircraft operational costs and turnaround efficiency at airports. *Journal of Air Transport Management*, *6*(4), 201–208. https://doi.org/10.1016/S0969-6997(00)00014-4

Xu, Y., Kang, M., Chen, D., Wu, C., Wang, X., Tu, R., & Yu, W. (2025). Flight departure optimization reduces airport CO2 emissions proved by Beijing-Tianjin-Hebei multiple-airport-region. *Transportation Research Part D: Transport and Environment*, *147*, 104920. https://doi.org/10.1016/j.trd.2025.104920

A
Literature Review Procedure and Selected Studies

The first literature search was conducted in Scopus using the query TITLE-ABS-KEY(("turnaround" OR "turn around") AND ("digital twin" OR "digital twins")). These terms were selected to identify studies specifically addressing the use of digital twin technology in aircraft turnaround operations. The search yielded 38 records. To improve the academic quality of the sample, conference papers, conference reviews, and notes were removed prior to screening, leaving 17 records for title and abstract review. These records were screened for relevance to aviation, turnaround processes, and digital twin implementation. As a result, 15 records were excluded because they were outside the scope of the study, mainly due to a focus on healthcare, non-aviation domains, or the absence of a concrete digital twin application in turnaround operations. The remaining 2 studies were retrieved in full text and assessed for eligibility. Both met the inclusion criteria and were therefore included in the review. The flow of this first search is shown in figure A.1, while an overview of the included studies is provided in table A.1.

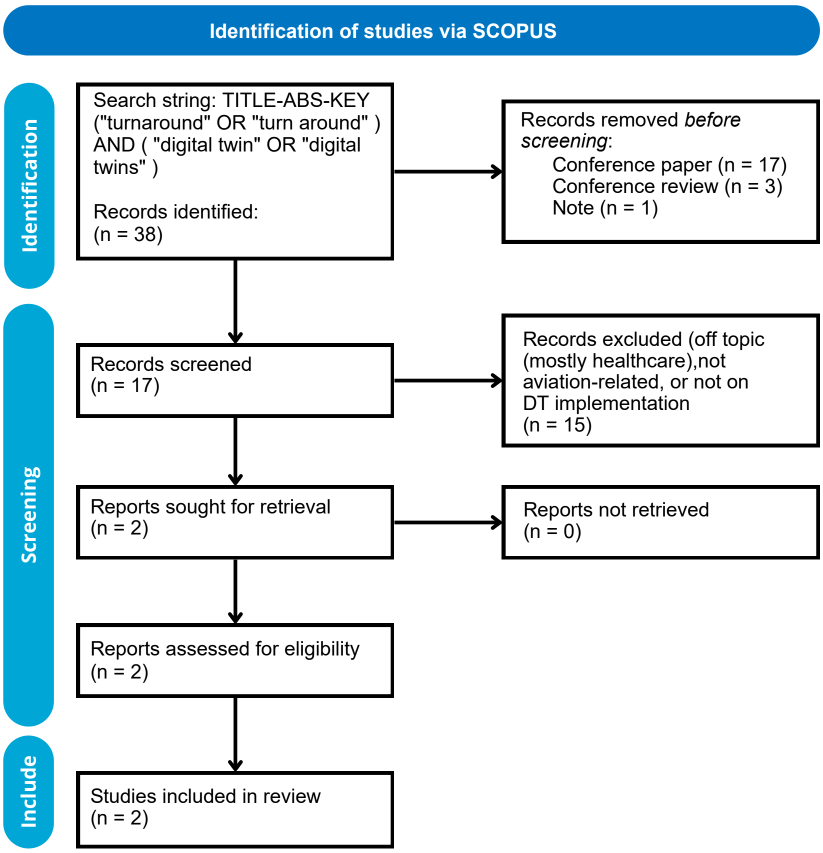


Figure A.1: Literature review flow "digital twin" AND turnaround.

Table A.1: Comparison of two turnaround-focused digital twin studies

Source	Study Aim	Digital twin Role	Case Context	Relevance
(Conde et al., 2022)	Improve turnaround management through event information exchange	Information-centric digital twin / interoperability architecture	Aberdeen International Airport	Shows digital twin as a means to improve real-time visibility, coordination, and data integration in turnaround management.
(Lu et al., 2025)	Predict and improve turnaround efficiency in an automated apron setting	Simulation- and control-oriented digital twin	Sandbox-based apron model using data from a Chinese airport	Shows digital twin as a tool for optimisation and automation, including predicted turnaround time reduction.

The second literature search was conducted in Scopus using the query TITLE-ABS-KEY("digital twin" AND airport). These terms were chosen to focus specifically on the application of digital twin technology within the entire airport environment rather than a single subsystem. The search produced 110 results. To ensure academic rigor, filters were applied to include only peer-reviewed journal articles, excluding conference papers, reviews, and book chapters, which reduced the results to 38 articles. Titles and abstracts were screened for relevance to digital twin implementation, airport operations, and socio-technical integration, narrowing the selection to 14 studies. Four could not be accessed in full text, leaving 10 key journal articles for analysis, and the flow of research is presented in figure A.2. An overview of these selected studies, including their focus areas and case contexts, is presented in table A.2.

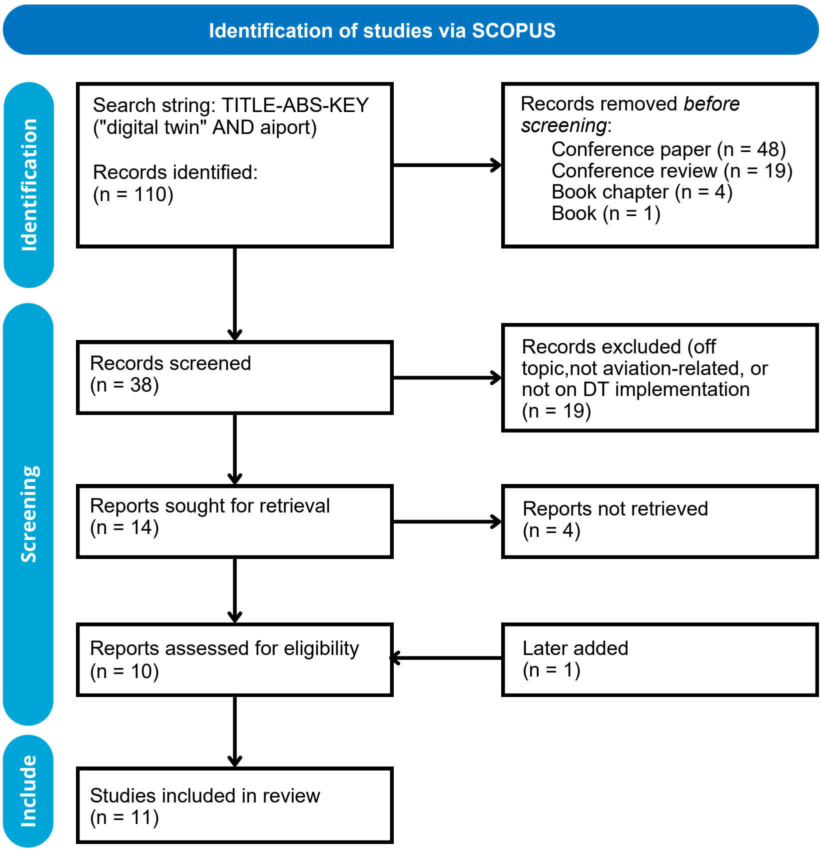


Figure A.2: Literature review flow "digital twin" AND airport.

Table A.2: Overview of selected digital twin and airport studies.

Source	Digital twin focus area	Core problem addressed	Case / data context
(Abdelalim et al., 2025)	Facilities management / AI-driven predictive maintenance	Reducing maintenance costs and energy use through integrated digital twin platforms.	Mega-facilities, including airports and hospitals
(Attar et al., 2024)	Airside operations / simulation-based optimization	Enhancing airside capacity and reducing delays through simulation-based digital twins.	International airport in Iran, using 2019 operational data
(K. Chen et al., 2024)	Air traffic management / human–system interaction	Improving collaborative aircraft routing through mixed-reality interfaces.	Singapore Changi Airport
(Conde et al., 2022)	Ground operations / turnaround information management	Reducing delays through real-time, standardized data sharing for turnaround events.	Aberdeen International Airport
(Luo et al., 2024)	Ground handling / vehicle scheduling	Minimizing vehicle travel and waiting times with robust digital twin-based rescheduling.	Domestic Chinese hub airport
(Mizrak et al., 2024)	Strategic sustainability planning	Prioritizing sustainability technologies, including energy systems, Internet of Things applications, and digital twins, in airport strategy.	Istanbul Airport
(Moon et al., 2025)	Facilities management / information strategy planning	Addressing fragmented maintenance practices and data standards through smart technology integration.	14 regional airports in South Korea
(Qiu, 2024)	Air terminal yards / resource optimization	Minimizing transport distance and improving yard efficiency through dynamic digital twin scheduling.	Hypothetical air terminal yard case
(Saifutdinov et al., 2022)	Ground traffic control / emulation testing	Testing centralized airport ground traffic control under critical scenarios.	Generic airport simulation model
(Xu et al., 2025)	Multi-airport region / flight scheduling optimization	Optimizing upstream flight scheduling in a multi-airport region; relevant as wider airport-system context rather than turnaround-specific digital twin development.	Beijing–Tianjin–Hebei multi-airport region: ZBAA, ZBTJ, and ZBSJ
(C. Liu et al., 2024)	Ground service processes / delay diagnosis and propagation	Diagnosing ground-service delays in real time and analysing how delays propagate through linked aircraft turnaround service processes using an L3-level digital twin framework.	Chongqing Jiangbei International Airport T3A Terminal, using real-world ground-service data
Later added			
(Durmaz & Göçmen, 2025)	Airport management / digital transformation	Examining how digital twin technology supports airport management through infrastructure monitoring, maintenance planning, operational simulation, scenario analysis, and predictive maintenance, while also identifying implementation challenges such as data compatibility, human resources, and cybersecurity.	Istanbul Airport case study, based on document analysis and semi-structured interviews

The third literature search focused on Digital Twin maturity assessment to provide a more complete overview of the field. The search was conducted in Scopus using the query TITLE-ABS-KEY(("digital twin" OR "DT") AND maturity). This yielded 96 records. To ensure academic rigor, filters were applied to remove conference papers (n = 35), conference reviews (n = 14), review articles (n = 6), and non-English publications (n = 7), resulting in 34 records for screening. Titles and abstracts were then assessed for relevance to maturity models, assessment frameworks, and evaluation approaches for Digital Twins, excluding off-topic studies (n = 15). Of the remaining papers, 19 full texts were sought, but 8 could not be retrieved. Ultimately, 11 studies were assessed for eligibility and included in the review, forming the basis for the maturity assessment analysis , and the flow of research is presented in figure A.3. An overview of these selected studies, including their focus areas and main focus, is presented in table A.3.

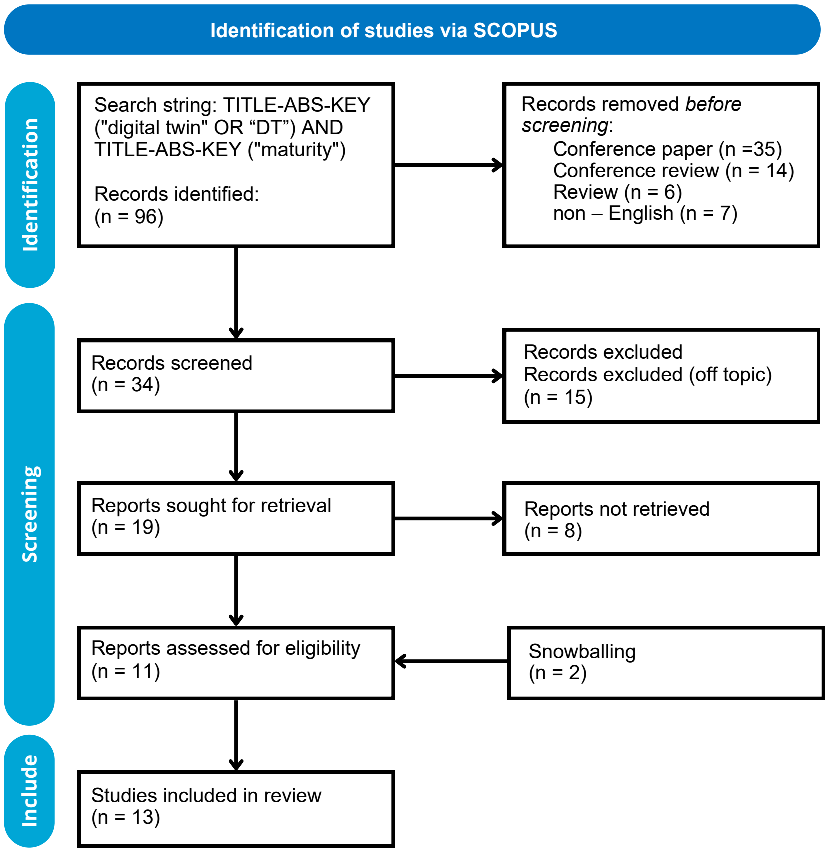


Figure A.3: Literature review flow "digital twin" AND maturity.

Table A.3: Overview of selected digital twin maturity model studies.

Paper	Research area (domain)	Main lens
(Haraguchi et al., 2024)	City digital twins / smart city planning	Governance-focused: links maturity to transparency, accountability, participation, and inclusion; also discusses risks such as exclusion and privacy/trust concerns (CITYSTEPS model).
(Hartmann & Wimmer, 2026)	Smart city digital twin maturity models (review/comparison)	Comparison-focused: reviews multiple digital twin maturity models across levels, capabilities, technologies, and data; points out gaps (e.g., the "full lifecycle" is often missing).
(Y. Liu et al., 2024)	Cross-domain digital twin maturity (review + proposed DTMM)	Roadmap-oriented: maturity levels are tied to capabilities and their technical requirements (and linked tools), to support planning and assessment.
(Manickam et al., 2023)	Digital twin maturity from a user experience perspective	User experience angle: maturity assessed through experience pillars (understandability, usability, convenience, dependability, delight) and how users explore, co-create, and decide with the twin.
(Scheibmeir & Malaiya, 2022)	Digital twins as discussed on social media (Twitter) + digital twin maturity model	Adoption/perception angle: uses tweet analysis (industries, topics such as Internet of Things and artificial intelligence) to inform a maturity model; useful for what the field talks about in practice.
(Uhlenkamp et al., 2022)	Production & logistics digital twins	Structured evaluation: DTMM with seven categories (e.g., context, data, computing, model, integration, control, human–machine interface) and 31 ranked characteristics; validated with use cases.
(Z.-S. Chen et al., 2024)	Building / construction digital twins	Quantitative assessment: maturity measured using multiobjective optimization and "fairness-aware" collective expert judgement; aimed at assessing real building digital twin projects.
(L. Chen et al., 2021)	Asset management digital twins (built environment)	Principles-driven: maturity dimensions/subdimensions/rubrics grounded in the Gemini Principles; uses expert input and case illustrations for project-level assessment and roadmapping.
(Li et al., 2024)	Underground infrastructure digital twins (tunnels/subways)	Quantitative MCDM: maturity hierarchy + rubrics with weighted evaluation using DEMATEL + ANP (D-ANP) to capture interdependencies; includes station/tunnel cases.
(Mo et al., 2023)	Autonomous manufacturing systems (Industry 4.0 context)	Autonomy maturity: not purely a digital twin maturity model; focuses on maturity of autonomy functions and enabling technologies (useful if digital twin maturity includes autonomy/self-optimizing elements).
(Gourisetti et al., 2025)	Digital twin technology in the energy sector (chapter)	Security/privacy + modeling: discusses digital twin modeling processes alongside cybersecurity and regulatory issues; useful for justifying security and privacy requirements in maturity work.
(Halpern et al., 2021)	Airport digital maturity (digital transformation, passenger experience)	Airport digital maturity (not digital twin-specific): proposes an airport digital maturity model and separates technological and organisational transformation; useful background for airport-sector maturity framing.
Snowballing		
(Madni et al., 2019)	Model-based systems engineering / digital twin integration	Foundational systems-engineering lens: explains how digital twins extend virtual prototypes by being continuously updated with performance, maintenance, and health data; links digital twins to IoT, simulation, digital thread, lifecycle traceability, what-if analysis, and decision support.
(Van Der Horn & Mahadevan, 2021)	Cross-domain digital twin definition, characterisation, and implementation	Conceptual and implementation lens: reviews digital twin definitions and proposes a generalised definition based on a virtual representation, a physical system, and information exchange; useful for distinguishing digital twins from other digital technologies and for defining core digital twin components and implementation considerations.

B
Literature-Based Maturity Assessment Framework Used for Expert Review

This appendix presents the complete literature-based version of the digital twin maturity assessment framework that was used as the basis for the expert review. This version was developed before the refinement phase and therefore represents the initial framework based on the literature review and the requirements derived in the main text.

The appendix includes the literature-based maturity levels, the dimensions and indicators, and the preliminary threshold logic for assigning maturity levels. These elements formed the interview and focus group material for the framework refinement phase and were later refined based on feedback from academic, consultancy, airline, and airport perspectives.

The purpose of this appendix is to make the initial framework transparent and traceable. By presenting the full literature-based framework separately, the main text can focus on the design choices and refinements, while this appendix provides the detailed version that was presented to experts for feedback.

Table B.1: Literature-based maturity levels of the airport turnaround digital twin

Level	Name	Description
Level 1	Digital model	The turnaround process is represented digitally, but the model is largely static or manually updated. It mainly supports documentation and basic visualisation.
Level 2	Digital shadow	Operational data flow automatically from the physical turnaround environment to the digital environment. The system reflects operations in near real time, but interaction remains largely one-way.
Level 3	Monitoring twin	The digital twin provides integrated monitoring of aircraft status, task progress, delays, and resource usage. The system supports operational awareness across multiple turnaround actors.
Level 4	Predictive twin	The twin supports diagnosis, forecasting, and scenario analysis. It helps identify emerging problems and estimate likely operational consequences.
Level 5	Optimising twin	The twin supports prescriptive decision-making through dynamic rescheduling, resource optimisation, and recommendation of operational responses to disturbances.

Table B.2: Literature-based D1 Data infrastructure indicators

Code	Indicator	Indicator definition
D1 Data infrastructure		

Continued on next page

Table B.2: Literature-based D1 Data infrastructure indicators continued

Code	Indicator	Indicator definition
D1.1	Automatic data capture	Assesses whether turnaround data are collected automatically from relevant operational sources rather than entered manually.
D1.2	Timely data updating	Assesses whether operational data are updated frequently enough to reflect the current turnaround situation.
D1.3	Data completeness	Assesses whether the digital twin includes sufficiently complete data on relevant turnaround activities, resources, and milestones.
D1.4	Data quality management	Assesses whether data quality issues, such as inconsistency, inaccuracy, or missingness, are identified and handled systematically.
D1.5	Integration of live and historical data	Assesses whether real-time and historical data are available together to support both current awareness and longer-term analysis.

Table B.3: Literature-based D2 System integration indicators

Code	Indicator	Indicator definition
D2 System integration		
D2.1	Multi-stakeholder data integration	Assesses whether data from multiple turnaround stakeholders are integrated into the digital twin environment.
D2.2	Shared operational view	Assesses whether different operational systems are combined into one shared and coherent operational view.
D2.3	Shared identifiers and interfaces	Assesses whether common identifiers, data structures, or interfaces are used to connect different systems and sources.
D2.4	Cross-boundary information exchange	Assesses whether information can be exchanged across organisational boundaries with limited manual re-entry or duplication.

Table B.4: Literature-based D3 Predictive and analytical capability indicators

Code	Indicator	Indicator definition
D3 Predictive and analytical capability		
D3.1	Early problem detection	Assesses whether the digital twin can identify emerging turnaround problems before they fully materialise.
D3.2	Impact estimation	Assesses whether the digital twin can estimate the likely operational effects of disruptions or deviations.
D3.3	Scenario analysis	Assesses whether the digital twin supports what-if analysis or exploration of alternative operational scenarios.
D3.4	Data-driven prediction	Assesses whether the digital twin uses current and historical data to generate predictions about turnaround performance.
D3.5	Causal analysis support	Assesses whether the analytical outputs help users distinguish likely causes of problems from visible symptoms.

Table B.5: Literature-based D4 Decision support and optimization indicators

Code	Indicator	Indicator definition
D4 Decision support and optimization		
D4.1	Response selection support	Assesses whether the digital twin supports the selection of operational responses when disruptions or deviations occur.
D4.2	Resource adjustment support	Assesses whether the digital twin supports the reassignment of staff, vehicles, stands, or other resources when conditions change.
D4.3	Task and schedule adjustment support	Assesses whether the digital twin supports changes in the order or timing of turnaround tasks in response to operational developments.
D4.4	Comparison of response options	Assesses whether alternative operational responses and their likely consequences can be compared systematically.
D4.5	Actionable recommendations	Assesses whether the digital twin provides recommendations or priorities that can be directly used in operational practice.

Table B.6: Literature-based D5 Operational representation indicators

Code	Indicator	Indicator definition
D5 Operational representation		
D5.1	Current turnaround status visibility	Assesses whether the digital twin clearly shows the current status of aircraft, turnaround tasks, and relevant resources.
D5.2	Task dependency visibility	Assesses whether the digital twin clearly shows how turnaround activities depend on and affect one another.
D5.3	Disruption visibility	Assesses whether delays, disruptions, and abnormal operational events are clearly represented.
D5.4	Operational level of detail	Assesses whether the representation contains enough operational detail to support understanding and decision-making.

Table B.7: Literature-based D6 Human interaction and collaboration indicators

Code	Indicator	Indicator definition
D6 Human interaction and collaboration		
D6.1	Understandability	Assesses whether the dashboard or interface is understandable for its intended users.
D6.2	Usability in practice	Assesses whether the interface can be used effectively in day-to-day turnaround work.
D6.3	Shared situational awareness	Assesses whether the interface supports a shared understanding of the turnaround situation across different users and actors.
D6.4	Information trustworthiness	Assesses whether users consider the presented information sufficiently trustworthy for operational interpretation and decision-making.
D6.5	Collaboration support	Assesses whether the interface supports collaboration and coordination between turnaround actors.

Table B.8: Literature-based D7 Governance, interoperability, and trust indicators

Code	Indicator	Indicator definition
D7 Governance, interoperability, and trust		
D7.1	Clarity of roles and responsibilities	Assesses whether roles, responsibilities, and ownership arrangements related to the digital twin are clearly defined.
D7.2	Formalised data-sharing rules	Assesses whether rules for data sharing, data use, and related governance procedures are formally established.
D7.3	Privacy, access control, and security	Assesses whether appropriate measures for access control, privacy protection, and data security are defined and embedded.
D7.4	Alignment with operational and accountability structures	Assesses whether the digital twin setup is aligned with airport operational rules, safety requirements, and accountability arrangements.
D7.5	Future interoperability	Assesses whether the digital twin is structured in a way that enables future integration with other airport systems and stakeholders.

Table B.9: Threshold logic for assigning maturity levels

Level	Label	Main capability	Key indicators and threshold
1	Digital model	Turnaround is represented digitally	No formal threshold; basic digital representation exists.
2	Digital shadow	Operational data automatically update the system	**D1.1** Automatic data capture ≥ 3; **D1.2** Timely data updating ≥ 3; **D1.3** Data completeness ≥ 2; **D2.1** Multi-stakeholder data integration ≥ 2
3	Monitoring twin	Shared monitoring across multiple actors	**D1.3** Data completeness ≥ 3; **D2.1** Multi-stakeholder data integration ≥ 3; **D2.2** Shared operational view ≥ 3; **D6.3** Shared situational awareness ≥ 3
4	Predictive twin	Forecasting and impact estimation	**D3.1** Early problem detection ≥ 3; **D3.2** Impact estimation ≥ 3; **D3.3** Data-driven prediction ≥ 3; **D3.4** Scenario analysis ≥ 3; **D3.5** Causal analysis support ≥ 3
5	Optimising twin	Decision support and dynamic optimisation	**D4.1** Response selection support ≥ 4; **D4.2** Resource adjustment support ≥ 4; **D4.3** Task and schedule adjustment support ≥ 4; **D4.4** Comparison of response options ≥ 4; **D4.5** Actionable recommendations ≥ 4

C
Expert Interview Materials for Literature-Based Framework Refinement

C.1. Interview Documentation

This appendix provides an overview of the expert interviews conducted during this research. The full interview transcripts are not included in the appendix in order to protect participant confidentiality and to keep the appendix focused on the research procedure rather than the raw data. Instead, this appendix presents the interview overview and the semi-structured topic list used to guide the interviews. The interview data were analysed thematically and used to refine the literature-based maturity assessment framework.

C.1.1. Overview of interviews

Table C.1 provides an overview of the interviews used in the refinement and contextualisation of the maturity assessment framework. Participant names are not included; instead, participants are referred to by anonymised codes.

Table C.1: Overview of interviews

Code	Participant(s)	Date	Duration	Interview focus
I1	Researcher	21 April 2026	1 hour and 6 minutes	Feedback on the literature-based maturity framework, including the maturity levels, dimensions, indicators, and preliminary scoring logic. (R)
I2	Consultant 1	21 April 2026	56 minutes and 55 seconds	Feedback on the maturity framework from a digital twin and consulting perspective, with emphasis on data foundations, practical implementation, and the use of digital twins in airport and infrastructure contexts. (C1)
I3	Schiphol focus group: Business Analyst, Process Owner, and Service Designer	22 April 2026	1 hour, 21 minutes and 29 seconds	Feedback on the operational applicability of the maturity framework in the Schiphol turnaround context, including Wilbur, data use, system integration, governance, trust, and practical assessment. (S1–S3)
I4	Two KLM participants: Operations Research Data Scientist and Data Scientist	1 May 2026	1 hour, 13 minutes and 21 seconds	Additional perspective on airport turnaround operations, airline-airport relations, data sharing, operational decision-making, and the implications of airline data constraints for digital twin maturity. (A1 and A2)
I5	Consultant 2	1 May 2026	1 hour, 03 minutes and 09 seconds	Feedback on the maturity framework from a broader aviation and implementation perspective, with emphasis on the distinction between Digital Shadow and Monitoring Twin, the role of automation, the separation of governance from internal maturity, the overlap between dimensions, and the practical use of the assessment in a workshop setting. (C2)

C.1.2. Semi-structured interview guide

The interviews followed a semi-structured format. The same general topic list was used across the interviews to ensure comparability, but the exact follow-up questions differed depending on the participant's expertise and the direction of the discussion. This allowed the interviews to remain focused on the maturity framework while still leaving room for participants to raise additional issues.

General interview topics

1. **Introduction and purpose of the interview**
   - The aim of the research was introduced: developing a maturity assessment framework for digital twins in airport turnaround operations.
   - The participant was informed that the interview was intended to support the design and refinement of the framework.
   - The participant was asked to provide critical feedback on unclear, overlapping, missing, or difficult-to-assess elements.
2. **Maturity levels**
   - The participant was asked whether the five proposed maturity levels were clear and logically ordered.
   - The participant was asked whether the progression from a digital model to a digital shadow, monitoring twin, predictive twin, and optimising twin matched practical understanding.

- The participant was asked whether the levels were sufficiently distinct from one another.
- The participant was asked whether the level descriptions were suitable for the airport turnaround context.

3. **Maturity dimensions**

   - The participant was asked whether the seven maturity dimensions were clear, distinct, and complete.
   - The participant was asked whether the dimensions captured the main areas that determine the practical usefulness of a turnaround digital twin.
   - The discussion covered the following dimensions:
     - data infrastructure;
     - system integration;
     - predictive and analytical capability;
     - decision support and optimisation;
     - operational representation;
     - human interaction and collaboration;
     - governance, interoperability, and trust.
   - The participant was asked whether any dimensions were missing, overlapping, or less relevant.

4. **Indicators**

   - The participant was asked whether the proposed indicators were understandable and clearly formulated.
   - The participant was asked whether the indicators were observable and specific enough to be scored consistently.
   - The participant was asked whether important indicators were missing.
   - The participant was asked whether any indicators should be renamed, merged, removed, or moved to another dimension.
   - The participant was asked whether some indicators would be difficult to assess in practice.

5. **Scoring and assessment logic**

   - The participant was asked whether the framework could be used to assess digital twin maturity in practice.
   - The participant was asked whether a scoring approach could meaningfully distinguish between different maturity levels.
   - The participant was asked whether some indicators should carry more weight than others.
   - The participant was asked whether some indicators should be treated as supporting information rather than as direct determinants of the maturity level.

6. **Results and application**

   - The participant was asked what type of output would be most useful in practice.
   - The discussion included possible outputs such as an overall maturity score, separate dimension scores, a visual maturity profile, or another form of result presentation.
   - The participant was asked whether the framework should mainly support comparison, improvement, or both.
   - The participant was asked whether the results should only show the current maturity state or also suggest priorities and next steps for development.
   - The participant was asked what would make the results clear and useful enough to support decision-making.

7. **Final reflections**

   - The participant was invited to provide any remaining comments on the framework.
   - The participant was asked whether any important aspects had not yet been discussed.
   - The participant was asked what changes would make the framework more understandable, useful, or practically applicable.

C.1.3. Use of interview data

The interview findings are reported in the thesis through anonymised participant codes. The interviews are referred to as I1, I2, I3, and I4. Individual participants are referred to by their role-based codes, such as R for the researcher, C1 for the consultant, S1–S3 for the Schiphol focus group participants, and A1–A2 for the KLM participants. The findings are mainly paraphrased rather than quoted directly. This was done to protect confidentiality and to focus on the analytical meaning of the interviews rather than on the exact wording of individual participants.

The full transcripts and interview notes were used during the analysis but are not reproduced in the thesis appendix. They are stored separately and were used only for the purpose of this research.

D
Follow-Up Expert Interview on the Refined Framework

This appendix documents the follow-up expert review interview on the refined maturity assessment framework. The interview was conducted after the literature-based framework had been revised based on the first round of expert feedback. The purpose of the interview was to check whether the refined framework was clear, internally consistent, and suitable for use in the demonstration phase.

Table D.1: Overview of follow-up expert review interview

Code	Participant(s)	Date	Duration	Interview focus
I6	Researcher 1	7 May 2026	54 minutes and 27 seconds	Follow-up feedback on the refined maturity model, including the distinction between internal maturity and external implementation constraints, the assessment order, the possible addition of a baseline level, and the visualisation of assessment results.

The follow-up interview was conducted with the same researcher who participated in Interview 1. While the first interview focused on the initial literature-based maturity levels, dimensions, indicators, and scoring logic, this follow-up interview focused on the revised structure of the framework. In particular, the interview was used to discuss whether external implementation constraints should be separated from the internal maturity assessment, whether a baseline level should be added, how the model could be applied in a workshop setting, and how the assessment results could be visualised.

Follow-up interview topics

1. **Changes made to the framework**
   - The revised maturity levels, dimensions, and indicators were presented.
   - The participant was asked whether the refinements made the model more coherent.
   - Particular attention was given to indicators that had been made more specific, combined, or removed because of overlap.
2. **Final maturity level**
   - The distinction between an optimising twin and a fully automated twin was discussed.
   - The participant was asked whether it was appropriate to describe the highest maturity level as an optimising twin rather than an automated twin.
   - The discussion focused on the specific nature of airport turnaround operations, where the digital twin primarily supports operational coordination and decision-making rather than directly performing physical tasks.
3. **External implementation constraints**
   - The revised distinction between external implementation constraints and internal maturity was discussed.

- The participant was asked whether it was appropriate to remove governance and regulatory aspects from the internal maturity dimensions and treat them instead as boundary conditions.
- The discussion focused on how regulation, data sharing, accountability, and operational rules can limit how mature a digital twin is allowed to become in practice.

4. **Internal maturity model**
   - The revised internal maturity model was presented.
   - The participant was asked whether the use of minimum maturity requirements for each level was understandable.
   - The discussion included the idea that not every dimension is required at every maturity level.
   - The participant was asked whether the structure appeared consistent and usable.
5. **Level zero**
   - The participant raised the possibility of adding a level zero.
   - Level zero was discussed as a way to represent situations where no digital twin capability is present.
   - The discussion also addressed whether zero scores should be shown explicitly rather than left blank in the assessment table.
6. **Assessment format and usability**
   - The practical use of the model in a workshop setting was discussed.
   - The participant was asked whether the indicators should be formulated as questions to make them easier for participants to interpret.
   - The discussion focused on the balance between keeping the model manageable and keeping it detailed enough to produce meaningful maturity results.
7. **Order of assessment**
   - The participant was asked whether the assessment should start with external implementation constraints or with the internal maturity model.
   - The discussion showed that both orders are possible, depending on the purpose of the assessment.
   - Starting with the external constraints was discussed as a way to set the boundaries of the assessment.
   - Starting with the internal maturity model was discussed as a way to first identify what the system is technically capable of, before identifying which external constraints limit its use.
8. **Visualisation of results**
   - The participant was asked for feedback on possible ways to visualise the maturity assessment results.
   - The discussion included tables, coloured maturity blocks, bar charts, spider diagrams, and dimension-level maturity profiles.
   - The participant noted that the table was useful for showing the requirements for each maturity level, while more graphical visualisations could support communication of the results.
9. **Final reflections**
   - The participant was invited to provide final comments on the refined model.
   - The discussion confirmed that the revised model was considered understandable and suitable for use in the upcoming workshop, with some small points for improvement.

The findings from this follow-up interview were used to further refine the final maturity assessment framework. In particular, the interview supported the distinction between external implementation constraints and internal maturity, the addition of an explicit level zero, the use of zero scores rather than blank cells, and the need to consider the purpose of the assessment when deciding whether to assess external constraints before or after internal maturity.

E
Detailed Scoring Descriptions for the Internal Maturity Indicators

This appendix provides the detailed 0–5 scoring descriptions for the internal maturity indicators of the final framework. These descriptions support the practical application of Component 3 of the framework: the internal maturity dimensions and indicators.

The main text presents a compact overview of the final dimensions and indicators, while this appendix provides the full scoring descriptions used to interpret each indicator during the assessment. A score of 0 indicates that the capability is not present or not assessable. Higher scores indicate more advanced and operationally usable digital twin capability, moving from basic representation towards monitoring, prediction, and optimisation-oriented decision support.

The scoring descriptions should be interpreted as a structured application proposal rather than as a fully validated quantitative measurement scale. Their purpose is to support consistent interpretation during the demonstration and future use of the framework.

E.1. D1 — Data

This dimension assesses whether the digital twin has the necessary data foundation to represent the turnaround process, monitor the current operation, predict future states, and inform operational decisions. The scoring descriptions for the indicators in this dimension are shown in Table E.1.

Table E.1: Final framework component 3a: scoring descriptions for D1 data

Score	Description
D1.1 Automatic data capture	
Assesses whether relevant turnaround data are captured automatically from operational sources.	
0	No relevant turnaround data are digitally captured.
1	Data are mainly entered manually or exist only as static digital data.
2	Some key turnaround data are captured automatically.
3	Key data categories within the defined turnaround scope are captured automatically.
4	Most critical data needed for monitoring and prediction are captured automatically.
5	Critical data needed for optimisation are captured automatically and continuously within the defined scope.
D1.2 Timely data availability	
Assesses whether captured data become available quickly enough for the intended operational use.	
0	Captured data are not available in the digital twin.
1	Data become available too late for operational use.
2	Selected key data are available in real time or near real time.
3	Key data are available quickly enough to monitor the current turnaround situation.

Continued on next page

Score	Description
4	Data are available quickly and reliably enough for prediction, risk detection, and impact estimation.
5	Data are available with the speed, reliability, and continuity needed for decision logic and workflow-ready actions.
D1.3 Data completeness	
Assesses whether the digital twin contains the core turnaround data needed for representation, prediction, and decision support.	
0	Core turnaround data are largely missing.
1	Only basic static data are available, such as schedule, aircraft type, airline, stand/gate, or planned turnaround time.
2	Some live operational data are available, but key categories are missing.
3	Most core turnaround data are available, such as flight status, task progress, resources, actors, and basic disruptions.
4	Core data are complete enough for prediction and impact estimation.
5	Core data are complete and detailed enough within the defined scope to enable optimisation and structured decision logic.
D1.4 Data quality management	
Assesses whether data quality issues are systematically identified and addressed.	
0	No systematic attention is given to data quality.
1	Data quality issues are handled manually or only when users notice them.
2	Basic quality checks exist for selected important data.
3	Missing, delayed, inaccurate, or inconsistent data are systematically identified and addressed.
4	Data quality is actively monitored because prediction and impact estimation require reliable data.
5	Data quality is continuously monitored, validated, and improved because the data are used for optimisation and workflow-ready decisions.
D1.5 Historical operational data	
Assesses whether historical turnaround data are stored and used for learning, prediction, and improvement.	
0	No historical turnaround data are stored or used.
1	Historical data are fragmented or only used for offline reflection.
2	Historical data are stored for selected turnaround activities, but not used structurally.
3	Historical data are used for reporting, monitoring, and operational learning.
4	Historical data are used to identify patterns, predict risks, and estimate operational impacts.
5	Historical data are used for optimisation, decision logic, continuous improvement, and learning from previous turnaround scenarios.

E.2. D2 — System integration

This dimension assesses whether the digital twin is technically connected to the relevant systems, stakeholders, and data structures needed to create a consistent operational view. The scoring descriptions for the indicators in this dimension are shown in Table E.2.

Table E.2: Final framework component 3b: scoring descriptions for D2 system integration

Score	Description
D2.1 Integration of direct turnaround stakeholder data	
Assesses whether data from direct turnaround stakeholders are connected to the digital twin environment.	
0	No direct turnaround stakeholder data are integrated.
1	Data from one direct turnaround stakeholder are connected.
2	Data from a limited number of direct turnaround stakeholders are connected.
3	Data from several key direct turnaround stakeholders are connected, but important sources are still missing.
4	Data from most direct turnaround stakeholders are integrated in a structured and reliable way.
5	Data from all relevant direct turnaround stakeholders within the defined scope are integrated continuously and consistently.
D2.2 Shared operational data layer	
Assesses whether operational data are integrated into one consistent data layer for the digital twin.	
0	No shared operational data layer exists.
1	Data from separate systems are manually brought together.
2	Some data are combined, but the data layer remains fragmented.
3	A shared data layer exists for core live turnaround data.
4	The shared data layer connects most relevant operational systems reliably.
5	One unified operational data layer provides consistent input for prediction, optimisation, and workflow-ready actions.
D2.3 Standardised identifiers and interfaces	
Assesses whether connected systems use shared identifiers, formats, and interfaces to ensure that different sources refer to the same operational objects.	
0	Systems use separate identifiers, formats, and interfaces.
1	Some identifiers or interfaces are aligned within individual systems.
2	A few identifiers or interfaces are shared across systems, but inconsistencies remain.
3	Core turnaround objects use shared identifiers and interfaces.
4	Most connected systems use consistent identifiers, formats, and interfaces.
5	All relevant connected systems use standardised identifiers, formats, and interfaces for reliable exchange and workflow-ready actions.

E.3. D3 — Predictive and analytical capability

This dimension assesses whether the digital twin can analyse current and historical data to detect problems, predict future states, estimate impacts, and explore alternative scenarios. The scoring descriptions for the indicators in this dimension are shown in Table E.3.

Table E.3: Final framework component 3c: scoring descriptions for D3 predictive and analytical capability

Score	Description
D3.1 Data-driven prediction	
Assesses whether the digital twin uses live and historical data to predict future turnaround performance.	
0	The digital twin does not generate predictions.
1	Users make predictions mainly from experience, outside the digital twin.

Continued on next page

Score	Description
2	The digital twin generates basic predictions from selected data.
3	The digital twin uses live data to predict future turnaround performance for key parts of the operation.
4	The digital twin combines live and historical data to predict future performance, risks, or delays.
5	The digital twin continuously updates predictions and provides input for optimisation and workflow-ready decision logic.
D3.2 Early problem detection	
Assesses whether the digital twin identifies emerging turnaround problems before they fully materialise.	
0	The digital twin does not identify emerging problems or risks.
1	Users notice problems mainly after they have already affected the turnaround.
2	The digital twin shows basic warning signs for selected problems.
3	The digital twin detects emerging problems in the current turnaround situation.
4	The digital twin detects problems early enough for users to respond before major impact occurs.
5	The digital twin continuously detects emerging risks and gives timely warnings for optimisation or decision logic.
D3.3 Impact estimation	
Assesses whether the digital twin estimates which parts of the operation are likely to be affected by disruptions, deviations, or changing conditions, and which are not.	
0	The digital twin does not estimate operational impact.
1	Users judge impact mainly from experience or separate information sources.
2	The digital twin shows simple expected effects for selected disruptions or deviations.
3	The digital twin estimates likely effects on turnaround performance and affected tasks or resources.
4	The digital twin estimates impacts across connected turnaround activities, including possible knock-on effects.
5	The digital twin continuously updates impact estimates and provides input for prioritising operational decisions.
D3.4 Scenario analysis	
Assesses whether the digital twin explores what-if scenarios and shows their expected operational effects.	
0	The digital twin does not analyse scenarios.
1	Scenarios are explored manually outside the digital twin.
2	The digital twin allows simple what-if analysis for selected situations.
3	The digital twin compares alternative scenarios and shows their expected effects.
4	The digital twin simulates scenarios using live and historical data and shows relevant trade-offs.
5	The digital twin continuously evaluates relevant alternatives and provides scenario results for optimisation or workflow-ready actions.

E.4. D4 — Operational decision optimisation

This dimension assesses whether the digital twin translates analytical insights into structured decision options, recommendations, or workflow-ready actions within the defined turnaround scope. The scoring descriptions for the indicators in this dimension are shown in Table E.4.

Table E.4: Final framework component 3d: scoring descriptions for D4 operational decision optimisation

Score	Description
D4.1 Data-based response selection	
Assesses whether operational responses are selected using digital twin insights rather than mainly user experience.	
0	No operational responses are selected using the digital twin.
1	Users select responses mainly from experience.
2	Users check digital twin information, but decisions are still mostly experience-based.
3	Users regularly use digital twin insights to choose responses.
4	The digital twin suggests or ranks response options based on analytical insights.
5	The digital twin selects the preferred response using agreed decision logic, with human confirmation or override.
D4.2 Resource reassignment	
Assesses whether the digital twin uses resource availability to identify or propose changes to turnaround-relevant resources.	
0	Resource availability is not visible or used for reassignment.
1	Basic resource availability is visible, but reassignment is decided manually.
2	The digital twin shows which resources may be affected by changing conditions.
3	The digital twin identifies resource gaps or conflicts.
4	The digital twin proposes specific resource reassignments or priorities.
5	The digital twin creates a resource reassignment action that can be confirmed, sent, or logged.
D4.3 Task and schedule adjustment	
Assesses whether the digital twin identifies or proposes changes to task order, timing, or assignment.	
0	The digital twin gives no information for task or schedule changes.
1	The digital twin shows task delays or schedule deviations.
2	The digital twin shows affected task order, timing, or dependencies.
3	The digital twin identifies which tasks or timings need adjustment.
4	The digital twin proposes specific changes to task order, timing, or assignment.
5	The digital twin creates a task or schedule adjustment action that can be confirmed, sent, or logged.
D4.4 Workflow-ready action	
Assesses whether the digital twin provides a concrete action that can be confirmed, sent, logged, or routed through the workflow.	
0	The digital twin gives no concrete action.
1	The digital twin gives alerts, but no clear action.
2	The digital twin gives broad advice, but it is not linked to a specific actor or workflow.
3	The digital twin gives a specific action for a relevant actor or role.
4	The digital twin gives a prioritised action with clear reasoning.
5	The action can be confirmed, sent, logged, or routed through the operational workflow.

E.5. D5 — Operational situation visibility

This dimension assesses whether the digital twin makes the current turnaround situation visible, understandable, and operationally meaningful for users. The scoring descriptions for the indicators in this dimension are shown in Table E.5.

Table E.5: Final framework component 3e: scoring descriptions for D5 operational situation visibility

Score	Description
D5.1 Current turnaround situation visibility	
Assesses whether the current status of aircraft, tasks, and resources is clearly represented.	
0	No current turnaround status is visible.
1	Only basic static information is visible.
2	Some current status information is visible, but it is incomplete or delayed.
3	The current status of aircraft, tasks, and resources is visible for key parts of the turnaround.
4	The current turnaround situation is clearly visualised across most relevant tasks and resources.
5	The full current turnaround situation is clearly visualised and immediately understandable for operational use.
D5.2 Task dependency visibility	
Assesses whether relationships and dependencies between turnaround tasks are clearly shown.	
0	Task dependencies are not visible.
1	Tasks are shown separately, without clear links.
2	Some basic task relationships are visible.
3	Key task dependencies are visible within the turnaround process.
4	Task dependencies are clearly shown, including how one task affects another.
5	Task dependencies are clearly shown and usable for task resequencing or schedule adjustment.
D5.3 Disruption visibility	
Assesses whether delays, disruptions, and abnormal events are clearly represented and easy to interpret.	
0	Delays or disruptions are not visible.
1	Delays or disruptions are only visible through manual interpretation.
2	Some delays or disruptions are shown, but not clearly.
3	Key delays and disruptions in the turnaround process are clearly visible.
4	Delays, disruptions, and abnormal events are clearly represented and easy to interpret.
5	Disruptions are immediately visible, easy to interpret, and clearly linked to affected tasks, resources, or turnaround steps.
D5.4 Role-relevant operational detail	
Assesses whether users are shown the operational detail relevant to their role without information overload.	
0	No useful operational detail is shown.
1	Very limited detail is shown, or the detail is not relevant for users.
2	Some relevant detail is shown, but users still need extra information elsewhere.
3	Relevant operational detail is shown for key users and common decisions.
4	The representation shows the right level of detail for most users and decisions.
5	Each user sees the relevant detail needed to understand the situation and act on their role, without information overload.

E.6. D6 — Human interaction and trust

This dimension assesses whether intended users can understand, use, and trust the digital twin in real turnaround operations. The scoring descriptions for the indicators in this dimension are shown in Table E.6.

Table E.6: Final framework component 3f: scoring descriptions for D6 human interaction and trust

Score	Description
D6.1 Understandability	
Assesses whether intended users can understand the interface and presented information without specialist data expertise.	
0	Users cannot understand the interface or information.
1	Users need specialist knowledge or extensive explanation.
2	Users understand basic information, but often need extra context.
3	Intended users understand the main turnaround situation without specialist explanation.
4	Intended users understand the situation, key changes, and predicted consequences.
5	Each intended user group quickly understands the information and recommendations relevant to its role.
D6.2 Usability in practice	
Assesses whether intended users can use the digital twin effectively in daily turnaround work.	
0	Users cannot use the digital twin in daily work.
1	Users can only use it with frequent help, training, or workarounds.
2	Users can perform basic tasks, but use is slow or inconsistent.
3	Users can use it for common daily turnaround tasks without major difficulty.
4	Users can use it effectively during time-sensitive operational work.
5	Use is smooth, fast, and embedded in daily workflows, including recommendation or action confirmation where relevant.
D6.3 Information trustworthiness	
Assesses whether users consider the digital twin's information reliable enough for operational interpretation and decision-making.	
0	Users do not trust the information.
1	Users rarely trust the information without checking other sources.
2	Users trust basic information, but often verify important details elsewhere.
3	Users generally trust the information for routine operational interpretation.
4	Users trust the information for time-sensitive interpretation and decision-making.
5	Users trust the information and recommendations enough to act on them, with checks only for exceptions or required procedures.

F
Schiphol Demonstration Documentation and Assessment Results

This appendix presents the scoring results from the Schiphol demonstration session. The results are divided into the external implementation constraints and the internal maturity assessment. The individual scores reflect the first interpretation of the participants, while the final score reflects the shared interpretation after group discussion.

Table F.1: Overview of framework demonstration session

Code	Participant(s)	Date	Duration	Session focus
I7	Schiphol demonstration group: Business Analyst, Process Owner, Service Designer, UX Designer, and Service Owner Aircraft	11 June 2026	1 hour, 58 minutes and 19 seconds	Demonstration and application of the final maturity assessment framework to the Schiphol turnaround digital twin context. The session focused on assessing the external implementation constraints and internal maturity dimensions, discussing the clarity of the scoring logic, interpreting the resulting maturity level, and evaluating the usefulness of the framework for identifying capability gaps and development priorities. (S1–S5)

F.0.1. External constraints Schiphol

Table F.2: External constraint scores for Schiphol

External constraint	Score	Notes
Data access, ownership and privacy	4	The group considered data access, ownership and privacy to be relatively mature. Participants explained that role-based access is used to give users access to pages, functionalities and insights that fit their role. At the same time, Wilbur is mainly a display layer and Schiphol is not always the owner of the source data. The score was therefore interpreted as level 4 rather than level 5, because not all desired data are available or complete.
Commercial and competition constraints	3–4	The discussion showed that much operational information is shared, such as arrival times, departure times, TOBT-related information and other contextual data. However, passenger information was discussed as a sensitive data category. Participants noted that such information is not shown to everyone and is filtered by role or organisation. The final interpretation was therefore between level 3 and level 4.
Decision, accountability and workflow	3–4	Participants indicated that the system is around level 3, with some parts moving towards level 4. Gate planning was mentioned as more advanced, while turnaround handling was seen as closer to level 3. The discussion also showed that clear rules for interpretation, accountability, escalation and traceability are not yet fully established.
Safety constraints	2	SSafety was identified as difficult to score. Participants indicated that Wilbur can display live information, but that it is not yet aligned strongly enough with all airside procedures to rely on it for safety-relevant operational decisions. The system still requires human checks, and participants noted that Wilbur is not mission-critical in the sense that the operation can rely on it without additional verification.

F.0.2. Internal maturity Schiphol

D1 Data

Table F.3: Scores for D1 Data

Ind.	S1	S2	S3	S4	Final	Notes
D1.1	5	4	4	4	4	Participants agreed on a final score of 4. Relevant data are captured automatically, but the group did not consider the coverage complete enough for level 5. Examples mentioned were that some data are not uniformly available, PACS data are not always available in the same way, and Deep Turn does not have full ramp coverage.
D1.2	4	5	5	5	5	The final score was 5. Participants indicated that data are available quickly enough and that the system does not require long waiting times for updates. One participant summarised this as speed, reliability and continuity.
D1.3	3–4, later 2	5, later 2	5 or 2	2–3	2	he final score was 2. The discussion showed that important data categories are still missing. Participants mentioned missing ramp coverage, missing asset status, no integration with handler resources, no integration with staff rosters or planning, and gaps in information about where towed aircraft are located. Participants also noted that this low score should be understood in relation to Schiphol's role, because the turnaround itself is largely performed by handlers.
D1.4	?	?	?	?	2	The final score was 2. Participants were uncertain about this indicator. Some checks and feedback mechanisms exist, and data quality issues may become visible through Wilbur or through other systems. However, participants also said that issues are not always detected proactively and that some quality checks are outside Wilbur itself. Therefore, the final interpretation was that basic checks exist, but data quality management is not yet systematic.
D1.5	3	3	4, later 3	3	3	The final score was 3. Historical data are used, but participants indicated that reporting, monitoring and operational learning often take place in other environments, such as Power BI or by analyst teams. Within the scope of Wilbur as the digital twin, historical data are not yet used in this composition for structural prediction, impact estimation or optimisation.

D2 System integration

Table F.4: Scores for D2 System integration

Ind.	S1	S2	S3	S4	Final	Notes
D2.1	3	3	3	3	3	All participants scored this indicator as 3. Several key data sources are connected, but important sources are still missing. The main missing element discussed was data from the handlers themselves, especially resource planning for people and equipment, such as push-back trucks.
D2.2	3	3–4	3–4	4	4	The final interpretation was 4. Participants indicated that the data are generally reliable and that much of the operational picture can be followed through the system. However, the data are not experienced as one fully unified layer. One participant noted that the data come from multiple layers, and another mentioned that blind spots remain in the turnaround picture.
D2.3	3	?	?	?	4	The final interpretation was 4. Participants discussed whether different systems use the same definitions, identifiers and interfaces. They indicated that the main operational objects are mostly aligned, for example through shared flight information, but that inconsistencies can still occur in the details or between

D3 Predictive and analytical capability

Table F.5: Scores for D3 Predictive and analytical capability

Ind.	S1	S2	S3	S4	Final	Notes
D3.1	5	4–5	4–5	4–5	4	The final score was 4. Participants recognised predictive capability, especially in relation to TOBT advice and Deep Turn. However, the group did not fully score this as 5 because the output is not yet workflow-ready decision logic. The advice indicates that something is expected to happen, but it does not yet provide a complete operational instruction.
D3.2	3	3–4	3–4	3	3	The final score was 3. Participants indicated that some emerging problems can be detected, for example through TOBT-related signals. However, the system does not always give a direct signal for what should happen next, such as what to do when a fuel truck is late.
D3.3	2	3	3	2	2	The final score was 2. Participants discussed that some simple effects are visible, such as a delay or TOBT advice. However, within the turnaround scope, the system does not yet systematically estimate knock-on effects across connected activities. Some higher-scoring examples were considered outside the selected turnaround scope.
D3.4	2	1–2	1	1–2	1	The final score was 1. Participants discussed whether a delay prediction could count as a simple what-if, but concluded that the system does not really compare scenarios. Scenario analysis mainly remains outside the digital twin itself.

D4 Operational decision optimisation

Table F.6: Scores for D4 Operational decision optimisation

Ind.	S1	S2	S3	S4	Final	Notes
D4.1	2–4	2	3	2	2	The final score was 2. Participants indicated that users may check digital twin information, but responses are still mostly based on operational experience and existing ways of working. The system does not generally present a concrete response option, except for specific examples such as weather-related measures.
D4.2	3–4	0–1	0–1	0	1	The final score was 1. Some resource-related information is visible, such as stands, ramps or equipment on the ramp page, but reassignment is not structurally supported. Participants emphasised that handler resources are not included and that reassignment remains separate from the digital twin.
D4.3	2	0	3	1	1	The final score was 1. The system may show delays or schedule deviations, but it does not operate at detailed task level. Participants noted that the system can advise a TOBT delay, but it does not adjust the order, timing or assignment of turnaround tasks.
D4.4	3	2	2	0	2	The final score was 2. The system can provide advice, such as delaying or advancing a TOBT, but this is not yet linked to a workflow in which the action can be directly confirmed, sent, logged or routed. Participants noted that Wilbur itself cannot yet make the adjustment.

D5 Operational situation visibility

Table F.7: Scores for D5 Operational situation visibility

Ind.	S1	S2	S3	S4	Final	Notes
D5.1	3	3	3	4	3	The final score was 3. The current situation is visible for key parts of the operation, but participants did not consider the visualisation clear enough across all roles and all relevant tasks and resources to justify level 4.
D5.2	2	1	1	1	1	The final score was 1. Tasks are mainly shown separately and dependencies within the turnaround are not clearly shown. A possible relation between maintenance and planned flights was discussed, but this was not considered a turnaround task dependency.
D5.3	3	3	3	3	3	The final score was 3. Key delays and disruptions are visible, but abnormal events are not fully represented. Participants gave the example that an event such as a flat tyre cannot be indicated in the system.
D5.4	2–3	2	?	3	2	The final score was 2. Some role-relevant operational detail is available, but participants indicated that users, especially handlers, still need information from their own resource planning systems or other sources.

D6 Human interaction and trust

Table F.8: Scores for D6 Human interaction and trust

Ind.	S1	S2	S3	S4	Final	Notes
D6.1	3–4	3–4	4	3–4	4	The final score was 4. Participants indicated that users generally understand what is shown in the system. However, one participant noted that understanding what is shown does not automatically mean that users act on it.
D6.2	4	4	4	4	4	The final score was 4. Participants agreed with this score. One participant noted that further improvement would require more time and more developers.
D6.3	3	3	4	3	3	The final score was 3. Participants indicated that when situations become time-sensitive or operationally critical, users often fall back on older ways of working, such as calling. CDM alerts were also mentioned as something that is not always actively acted upon.

G
YVR Transferability Evaluation Materials

This appendix documents the evaluation interview with a participant from Vancouver International Airport (YVR). The interview was used to evaluate whether the maturity assessment framework could also be understood and applied outside the Schiphol case. The full transcript is not included in the appendix in order to protect confidentiality and to keep the appendix focused on the research procedure and the use of the interview data.

Table G.1: Overview of YVR transferability evaluation interview

Code	Participant(s)	Date	Duration	Interview focus
I8	YVR airport flow manager (Y)	19 May 2026	1 hour, 21 minutes and 54 seconds	Evaluation of whether the maturity assessment framework can also be applied outside the Schiphol context, with attention to YVR's operational context, stakeholder structure, data-sharing arrangements, turnaround processes, and external implementation constraints.

The participant was selected because of their operational role at YVR, where they monitor airport flows from landside to airside, including passenger processes, baggage processes, gate planning, and operational disruptions. This made the interview relevant for assessing whether the framework could also be understood and applied in an airport with a different operational and governance context.

The interview focused on the applicability of the framework to YVR's turnaround context, the distinction between internal maturity and external implementation constraints, and the extent to which the framework would need adaptation for airports with different stakeholder structures, data-sharing arrangements, and operational responsibilities.

YVR evaluation interview topics

1. **Participant role and operational context**
   - The participant was asked to describe their role at YVR.
   - The discussion covered YVR's monitoring of passenger flows, baggage processes, gate planning, airside processes, and operational bottlenecks.
   - The participant explained that YVR's role in turnaround operations is mainly related to monitoring, gate planning, and coordination rather than direct control of ground handling activities.
2. **Current role of YVR in turnaround operations**
   - The participant was asked how YVR is currently involved in aircraft turnaround operations.
   - The discussion focused on YVR's current dependence on human monitoring, CCTV, gate planning, and communication with airlines and ground handlers.
   - The participant described YVR as being between a digital shadow and a monitoring twin for turnaround operations.
3. **Data sharing and stakeholder dependency**

- The participant was asked to what extent YVR depends on airlines, ground handlers, and other partners for turnaround data, decisions, and execution.
- The discussion covered the limited availability of airline and ground-handler data.
- Particular attention was given to passenger data, connection data, baggage data, load factors, and the commercial sensitivity of airline information.
- The participant explained that YVR is beginning to receive some airline data, but that this is still limited and not always available in real time.

4. **Internal maturity versus external implementation constraints**
   - The participant was asked whether the distinction between internal maturity and external implementation constraints was meaningful for YVR.
   - The discussion showed that YVR may develop internal technological capabilities faster than the external stakeholder agreements needed to fully use them.
   - The participant confirmed that collaboration with airlines and ground handlers is necessary before higher maturity levels can be reached in practice.
5. **Relevance of the maturity dimensions**
   - The participant was asked whether the six internal maturity dimensions were relevant for YVR.
   - The discussion covered data infrastructure, system integration, predictive and analytical capability, operational decision optimisation, operational visibility, and human interaction and trust.
   - The participant confirmed that the dimensions were relevant, but also noted that some dimensions become more relevant only at later maturity stages.
6. **External implementation constraints**
   - The participant was asked whether the external implementation constraints in the framework also apply to YVR.
   - The discussion covered data access, data ownership, privacy, commercial sensitivity, competition between airlines, decision accountability, workflow responsibilities, and safety.
   - The participant confirmed that these constraints are also relevant at YVR, although their specific form depends on the local regulatory and operational context.
7. **Safety, reliability, and trust**
   - The participant was asked about the role of safety in digital twin maturity.
   - The discussion showed that safety-related digital twin functions require high reliability, availability, and trust.
   - The participant emphasised that safety-critical digital twin functions would require fallback options and a high level of accuracy before they could be relied upon operationally.
8. **Transferability of the framework**
   - The participant was asked whether the framework is transferable outside Schiphol.
   - The participant concluded that the framework is transferable with adaptations.
   - The core structure of the framework was considered applicable, but the participant emphasised that each airport has its own operational, regulatory, and stakeholder context.
   - The discussion therefore supported the idea that the framework can be used outside Schiphol, but should not be treated as a one-size-fits-all assessment tool.

Use of the interview data

The findings from this interview were used to evaluate the transferability of the maturity assessment framework. The interview showed that the general structure of the model was understandable to an airport stakeholder outside Schiphol and that the core distinction between internal maturity and external implementation constraints was relevant in the YVR context. At the same time, the interview indicated that application of the framework requires adaptation to the local airport context, especially regarding data-sharing agreements, airline-airport relationships, operational responsibilities, safety requirements, and regulatory conditions.

The YVR evaluation therefore supports the conclusion that the framework is not limited to Schiphol, but that it should be applied as a flexible assessment framework. Its maturity levels, internal dimensions, and external constraints can guide assessment at other airports, provided that local operational and governance differences are taken into account.

H
Schiphol Post-Demonstration Evaluation Materials

This appendix presents the post-demonstration questionnaire that was sent to the Schiphol participants after the demonstration of the Digital Twin Maturity Assessment Framework. The questionnaire was used to evaluate the framework itself. It did not ask participants to reassess Schiphol's digital twin maturity, but focused on whether the framework was clear, useful, complete, and applicable for assessing digital twin maturity in airport turnaround operations.

The questionnaire addressed the maturity levels, dimensions, indicators, scoring logic, and the distinction between internal maturity and external implementation constraints. Because the questionnaire was only partly completed, the responses were not analysed as a complete survey dataset. Instead, the completed answers were used as supplementary qualitative evidence in the evaluation chapter.

H.0.1. Introduction shown to participants

Thank you for participating in the demonstration of the Digital Twin Maturity Assessment Framework.

This short questionnaire evaluates the framework itself. It does not ask you to reassess Schiphol's current maturity. Instead, it asks whether the framework was clear, useful, and complete enough to assess digital twin maturity in airport turnaround operations.

In this questionnaire, "the framework" refers to the maturity levels, dimensions, indicators, scoring logic, and the distinction between internal maturity and external implementation constraints.

Because this evaluation involves a small expert group, your explanation is more important than the category you select.

H.0.2. Response options

For each requirement, participants were asked to choose one of the following options.

Table H.1: Response options used in the Schiphol post-demonstration questionnaire

Option	Meaning
Fulfilled	The framework covers this well.
Partly fulfilled	The framework covers this, but improvement is needed.
Not fulfilled	The framework does not cover this well enough.
Unsure	I cannot judge this.

H.0.3. Participant background

P1. What is your role or job title?

Choose: (S1) Business Analyst, (S2) Process Owner, (S3) Service Designer, (S4) UX Designer, or (S5) Service Owner Aircraft.

Explain shortly.

P2. How are you involved in airport turnaround operations, airport digitalisation, digital twins, or operational decision-making?

Explain shortly.

H.1. Part A: Overall evaluation

A1. Did the framework help you discuss digital twin maturity in a structured way?

Please explain your answer.

A2. Did the framework help identify useful next development steps for digital twin development in turnaround operations?

Please explain your answer.

A3. Which part of the framework was most useful during the session?

A4. Which part of the framework was least useful or most difficult to apply?

H.2. Part B: Functional requirements

These questions evaluated whether the framework covered the digital twin capabilities that are relevant for turnaround operations. For each requirement, participants were asked to select one response option and provide a short explanation.

Table H.2: Functional requirement questions in the Schiphol post-demonstration questionnaire

Requirement	Evaluation question	Response option	Explanation
RF1. Real-time operational visibility	Does the framework clearly cover whether a digital twin can show the current turnaround situation, including status, resources, and disruptions?	Fulfilled / Partly fulfilled / Not fulfilled / Unsure	
RF2. Integrated and consistent data foundation	Does the framework clearly cover whether data from different systems and stakeholders can be combined into one consistent operational view?	Fulfilled / Partly fulfilled / Not fulfilled / Unsure	
RF3. Proactive operational intelligence	Does the framework clearly cover whether a digital twin can detect upcoming problems, identify likely causes, and estimate consequences?	Fulfilled / Partly fulfilled / Not fulfilled / Unsure	
RF4. Decision support for interventions	Does the framework clearly cover whether a digital twin can help users choose and compare possible responses to disruptions?	Fulfilled / Partly fulfilled / Not fulfilled / Unsure	
RF5. Resource optimisation	Does the framework clearly cover whether a digital twin can help improve the use of staff, vehicles, equipment, or stands?	Fulfilled / Partly fulfilled / Not fulfilled / Unsure	
RF6. Dynamic rescheduling	Does the framework clearly cover whether a digital twin can support changes to schedules, tasks, or resources when the situation changes?	Fulfilled / Partly fulfilled / Not fulfilled / Unsure	

Requirement	Evaluation question	Response option	Explanation
RF7. Shared situational awareness	Does the framework clearly cover whether different stakeholders can work from the same up-to-date operational picture?	Fulfilled / Partly fulfilled / Not fulfilled / Unsure	

B1. Which functional requirement is best covered by the framework? Why?

B2. Which functional requirement is least well covered or missing by the framework? Why?

H.3. Part C: Structural requirements

These questions evaluated whether the framework itself was organised in a clear and useful way. For each requirement, participants were asked to select one response option and provide a short explanation.

Table H.3: Structural requirement questions in the Schiphol post-demonstration questionnaire

Requirement	Evaluation question	Response option	Explanation
RS1. Multi-level maturity structure	Are the maturity levels clear and logical as a development path from a basic digital model to an optimising twin?	Fulfilled / Partly fulfilled / Not fulfilled / Unsure	
RS2. Multi-dimensional assessment	Is it useful that the framework assesses several dimensions instead of only giving one overall maturity score?	Fulfilled / Partly fulfilled / Not fulfilled / Unsure	
RS3. Modularity	Does the framework make it easy to see strengths and weaknesses in separate areas, such as data, integration, prediction, decision support, and usability?	Fulfilled / Partly fulfilled / Not fulfilled / Unsure	
RS4. Scoring clarity	Are the indicators and scoring descriptions clear enough to apply during an assessment?	Fulfilled / Partly fulfilled / Not fulfilled / Unsure	
RS5. Technical and organisational balance	Does the framework include both technical capabilities and organisational conditions, such as governance, workflows, and stakeholder collaboration?	Fulfilled / Partly fulfilled / Not fulfilled / Unsure	
RS6. Actionable development pathway	Does the framework help identify concrete gaps, priorities, or next steps for further digital twin development?	Fulfilled / Partly fulfilled / Not fulfilled / Unsure	

C1. Were the maturity levels and scoring descriptions clear enough? If not, what was unclear?

C2. Did the framework feel too detailed, too simple, or about right? Please explain.

H.4. Part D: Environmental requirements

These questions evaluated whether the framework fitted the real airport context. For each requirement, participants were asked to select one response option and provide a short explanation.

Table H.4: Environmental requirement questions in the Schiphol post-demonstration questionnaire

Requirement	Evaluation question	Response option	Explanation
RE1. Usability	Does the framework consider whether the digital twin can actually be used by operational and strategic users?	Fulfilled / Partly fulfilled / Not fulfilled / Unsure	
RE2. Understandability	Does the framework consider whether users can understand the information, indicators, and recommendations provided by the digital twin?	Fulfilled / Partly fulfilled / Not fulfilled / Unsure	
RE3. Interoperability	Does the framework consider whether information can be exchanged across systems, organisations, and stakeholders?	Fulfilled / Partly fulfilled / Not fulfilled / Unsure	
RE4. Governance compatibility	Does the framework consider airport constraints such as safety, accountability, regulation, decision rights, and operational procedures?	Fulfilled / Partly fulfilled / Not fulfilled / Unsure	
RE5. Data security and confidentiality	Does the framework consider whether operational data can be handled securely and confidentially?	Fulfilled / Partly fulfilled / Not fulfilled / Unsure	
RE6. Data and model trustworthiness	Does the framework consider whether the data, analytics, and digital twin outputs are reliable and trustworthy enough to support decisions?	Fulfilled / Partly fulfilled / Not fulfilled / Unsure	

D1. Did the distinction between internal maturity and external implementation constraints make sense?

Please explain why or why not.

D2. Are any important airport-specific constraints missing from the framework?

I
AI-statement

During the preparation of this thesis, I used NotebookLM (Google), ChatGPT (OpenAI), and the transcription function in Microsoft Word as support tools.

NotebookLM was mainly used to support the literature review process. It helped me navigate my own uploaded articles, identify relevant sections, compare article content, and check whether I had correctly interpreted specific parts of the literature. The original articles remained the primary sources, and all AI-supported outputs were checked against the source texts.

ChatGPT was used for brainstorming, structuring arguments, and refining written text. This included improving clarity, coherence, wording, and flow. ChatGPT was not used to generate original research findings, conduct the analysis independently, or draw final conclusions.

The transcription function in Microsoft Word was used to create initial transcripts of interviews and demonstration sessions. These transcripts were manually checked, corrected, and refined before being used in the research.

After using these tools, I critically reviewed, edited, and validated all content. I take full responsibility for the accuracy, originality, and integrity of the content of this thesis.