CMDB: Data Model Plan for Suomen Erillisverkot Oy

Sebastian Fallström

Degree Thesis

Information Technology

2023

Lärdomsprov

Sebastian Fallström

CMDB: Datamodellsplan för Suomen Erillisverkot oy

Yrkeshögskolan Arcada: Informationsteknik, 2023

Identifikationsnummer:

9306

Uppdragsgivare:

Suomen Erillisverkot oy

Sammandrag:

Studien skapar en möjlig datamodellsplan för Erillisverkot datacenterinfrastruktur till deras kommande Configuration Management Database (CMDB). Erillisverkot är ett företag som specialiserar sig i säkra ICT-lösningar. Huvudmålet med studien är att hantera utmaningarna inom IT-infrastrukturen genom att föreslå tre datamodeller som baserar sig på Configuration Items (CIs). Dessa inkluderar datacenter, rum, serverställ och hårdvara som ägs och används av företaget. Genom tidigare forskning, semistrukturerade intervjuer, observationer och workshopar i samarbete med konsultföretaget Justin presenteras en lösning. Datamodellerna för planen inkluderar två modeller som baserar sig på Justins Praktiska Datamodell för en CMDB (PDMC) modell, den högsta logiska datamodellen och den logiska datamodellen. Den tredje modellen som är den fysiska datamodellen, visar det sista steget före implementeringen. Modellerna har gjorts för att vara anpassningsbara, effektiva, lättlästa, och fungera med Erillisverkot nuvarande Data Center Infrastructure Management (DCIM) Netbox. Även om dessa modeller är konceptuella och fortfarande behöver testas före implementering, ger de en vägledning för planeringen av olika CMDB-datamodeller. Studien kan tjäna som en undersökning för organisationer som står inför liknande utmaningar.

Nyckelord:

CMDB, Datamodell, IT-infrastruktur, Konfigurationsobjekt, IT-tjänsthantering, Netbox DCIM, Erillisverkot.

Degree Thesis

Sebastian Fallström

CMDB: Data Model Plan for Suomen Erillisverkot Oy

Arcada University of Applied Sciences: Information Technology, 2023

Identification number:

9306

Commissioned by:

Suomen Erillisverkot Oy

Abstract:

The thesis focuses on creating a potential data model plan for the Configuration Management Database (CMDB) at Erillisverkot's data center infrastructure, a company that specializes in secure ICT solutions. The main goal of the study is to address the challenges of managing IT infrastructures by proposing a data model for the database to handle Configuration Items (CIs). These CIs include data centers, rooms, racks, and hardware owned and used by the company. Through previous research, semi-structured interviews, observations, and workshops in collaboration with the consultancy firm Justin, a solution is presented on three different levels. The thesis introduces three data models for CMDB implementation. The first two, the Highest Logical Data Model and the Logical Data Model, are based on the Practical Data Model for CMDB (PDMC), made by Justin Group. The third model, the Physical Data Model, represents the final step in the modeling process before the implementation. These models have been made to be adaptable, efficient, easy to read, and aligned with Erillisverkot's existing system, Netbox a Data Center Infrastructure Management (DCIM). Although these models are conceptual and still will need to be tested through implementation, they provide a good guide for planning CMDB data models. This study can serve as a study for organizations facing similar challenges.

Keywords:

CMDB, Data Model, IT Infrastructure, Configuration Items, IT Service Management, Netbox DCIM, Erillisverkot.

Contents

Abbreviations:

API – Application Programming Interface

DC – Data Center

CFS – Customer Facing Service

CDM – Conceptual Data Model

CIs – Configuration Items

CMDB – Configuration Management Database

CMS – Configuration Management System

CRM – Customer Relationship Management

CSDM – Common Service Data Model

DCIM – Data Center Infrastructure Management

ESP – External Service Provider

ICT – Information and Communication Technology

ISP – Internal Service Provider

ISO/IEC – IT service management

IT – Information Technology

ITIL – Information Technology Infrastructure Library

ITSM – IT Service Management

ISMS – Information Security Management Systems

LDM – Logical Data Model

PDM – Physical Data Model

PDMC – Practical Data Model for CMDB

SKMS – Service Knowledge Management System

1 Introduction

The digital era has brought numerous technical breakthroughs resulting in complex Information Technology (IT) infrastructures that require careful management to maintain smooth operation and service delivery. In this context, the Configuration Management Database (CMDB) is essential for competent IT Service Management (ITSM). The main focus of this thesis is to explore and design a potential planning phase, especially for the data model of the upcoming CMDB at Erillisverkot. A CMDB repository serves as a storage for relationships between information. This information, commonly called Configuration Items (CIs), includes hardware, software, and networks used by the company. The objective is to create a structured approach towards developing a data model that can effectively handle the complexities of the IT environment, ensuring effective representation and management of these CIs and their relationships. This research aims to provide potential strategies and insights to assist the company in creating a potential data model for the CMDB rather than offering a definitive solution for the entire CMDB system.

1.1 Background and Context

Erillisverkot is a state-owned Finnish company that provides secure ICT (Information and Communication Technologies) solutions to authorities and critical operators responsible for Finland's security. They offer Virve services, situational awareness tools, secure data centers, and cloud and cable location services. The company prioritizes security, reliability, and custom-tailored services for the customers.

At the start of the project, Erillisverkot enlisted the help of consultancy firm Justin (Justin Group Oy) to provide them with essential background information and to provide a guide to get them through the planning process. This thesis combines the knowledge gained from interviews, meetings, and Justin's real-world expertise to develop a potential data model for the company's upcoming CMDB.

1.2 Problem Statement

As Erillisverkot has grown, the lack of a centralized database has scattered information and data across numerous places, and a centralized repository for essential IT service-related components was needed. This kind of company structure resulted in ineffective information management and retravel, making everyone’s tasks harder. The demand for a single platform where workers from every department within the company can access, manage, and update data makes the CMDB an unbeatable database. The CMDB will eventually improve essential fields like communication, decision-making, and operational efficiency.

1.3 Objective of the study

The goal of this study is to provide Erillisverkot with an adaptable and actionable data model plan for the data center infrastructure. This study will use a combination of Justin’s material, insights from semi-structured interviews, observation from meetings, academic research, and the operational needs of Erillisverkot to develop a data model for their CMDB. This study will serve as a foundational step toward Erillisverkot implementing a CMDB system.

2 Literature Review

2.1 CMDB Overview

2.1.1 Introduction to CMDB

The Configuration Management Database (CMDB) is a key concept in the Information Technology Infrastructure Library (ITIL) framework, essential for IT Service Management (ITSM). It stores details about essential IT service elements, called Configuration Items (CIs). These CIs include a vast array of different components such as software, hardware, personnel, and even process paperwork (Drogseth, Sturm & Twing, 2015), “as well as the relationships between CIs and between CIs and IT services” (Brenner and Gillmeister (2014).

Both the ITIL framework and the ISO/IEC 20000 highlight the crucial role of CMDB in IT processes (Device42, 2023). The ITIL introduced the Service Knowledge Management System (SKMS), which is a bigger version of the CMDB. The SKMS collects data over an IT service's lifecycle to reflect its broader service management vision (Drogseth, Sturm & Twing, 2015). Inside the SKMS, the Configuration Management System (CMS) manages configuration data, with the CMDB as a specialized database within the CMS, focusing on storing the CIs and mapping the relationships (Brahmachary, 2018)

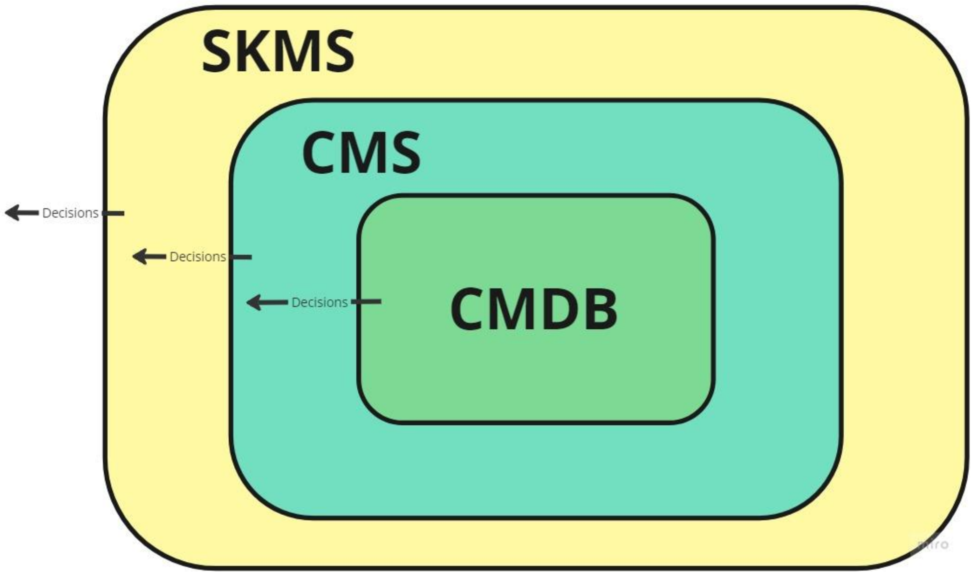

Figure 1, SKMS, CMS, and CMDB Hierarchical Structure (Adapted from Brahmachary, 2018)

2.1.2 Modern Implications of CMDB

The CMDB has significantly evolved with the challenges of today's companies. As more and more businesses are adopting cloud services, virtualization, and distributed systems, the importance of CMDB has become more crucial. Today, CMDB is not just an information database, but it is essential for overseeing complex IT ecosystems and ensuring that everything runs smoothly. (Device42, 2023).

However, modern CMDBs come with a set of challenges. Keeping them up-to-date and accurate with rapid technological changes is a difficult task for any company. However, by adopting best practices, as suggested by Device42, companies may use CMDB to turn obstacles into chances for improved ITSM.

2.1.3 Benefits of CMDB

The integration of a CMDB, which includes effective data models, can offer the following advantages, according to Simplilearn (2023):

- *Clerical Errors, Process Errors, and Programming Errors will be reduced*
- *Brakes down the barriers between IT and the business*
- *Provides more proactive management*
- *Help better assess risk, improve security*
- *Help keep track of any change in software*

2.2 Data Models and Their Role in CMDB

2.2.1 Understanding Data Models

A data model in the concept of the CMDB organizes and shows how different parts of the IT environment are related and work together, with the help of different CIs, their attributes, and the relationships between them. The data model contains tools and techniques to understand, document, and communicate how an organization should collect, update, store, and share data.

A well-structured data model is the central part of a successful implementation and operational CMDB. Such a model ensures clear and understandable data and leads to a simplified database design, better application development, and more accessible data retrieval. This structured approach is important for effective communication across different departments and helps stakeholders manage dependencies, changes, issues, risks, and decision-making processes. However, with an effective data model, the CMDB will become more complex and time-consuming, a concern highlighted in the literature by Brenner and Gillmeister (2014), underscoring that a well-designed data model is essential for grasping the complexity of the IT landscape.

With a diverse selection of CMDBs, there are specific data modeling standards and practices that stand out:

- Common Service Data Model (CSDM) by ServiceNow: "The common services data model can act as a blueprint to map your IT service on the ServiceNow platform" (ServiceNow, 2023b).
- Practical data model for CMDB (PDMC) made by Justin Group (2020b) is a perfect example of dividing the practical model into stages or levels, as they call them.

2.2.2 Data Model Components

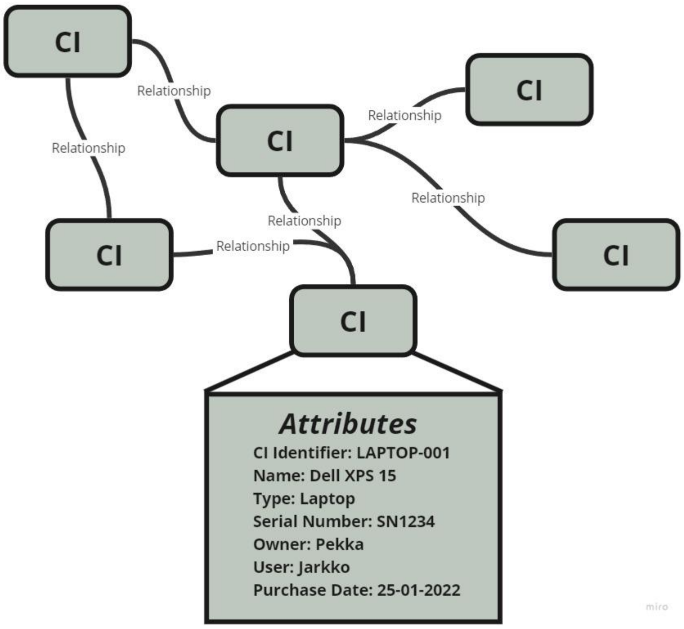

Figure 2, A simplified CMDB structure

At the core of a CMDB are the CIs, which are the primary building blocks. As illustrated in Figure 2, the company can add items as CIs, such as software, hardware, network components, offered services, documentation, people, and business processes. Each CI has specific attributes, which provide information about them, their status, and their role in the company's infrastructure.

The attributes are different kinds of information inside of the CI. The information given for every CI can vary depending on the category. For example, as shown in Figure 2, attributes include identification code, name or label, description, type or category, status, version, ownership, importance, vendor, and manufacturer. The goal of the fields is to track enough information to effectively manage and control the different CIs without overwhelming the team with unnecessary information.

The relationships between two CIs often comprise a Source CI and a related Target CI. These relationships define how one CI interacts with another CI. As outlined by ServiceNow (2023a), there are different types of relationships between the CIs:

- Depends on/Used by: This is one of the typical relationship types. This type is used, for example, between an application (CI) that depends on a server (CI). The connection will enable troubleshooting between the CIs if the server fails. (The connection enables us to see the failure of the application if the server fails.
- Runs on/Hosted by: Typically seen between software and hardware CIs. The application (CI) runs on the server (CI) or a virtual machine (CI).
- Parts of/Contains: Parts and larger systems or subsystems have these relationships. For example, a hard drive (CI) is a part of the server (CI).
- Connected to: Used between devices in racks, often network devices. For example, a switch (CI) inside a rack connected to a server (CI) inside the same or a different rack.
- Serves/Supported by: This is usually a relationship between services and the CIs that support it. For example, a service (CI) for sale relies on several software applications (CIs), which in turn depend on various hardware devices (CI).

The relationship will ensure that when one CI gets affected, will all the dependent or related CIs also be affected.

2.2.3 Types of Data Models

While managing the different complexities is essential, it is equally crucial to understand the different types of data models used in CMDBs. “Data modeling is a process that helps you to understand the structures, form, and relationships of information” Groves (2022). The planning process of data models typically includes three sequential phases, as outlined by Groves (2022): the Conceptual Data Model (CDM), the Logical Data Model (LDM), and the Physical Data Model (PDM).

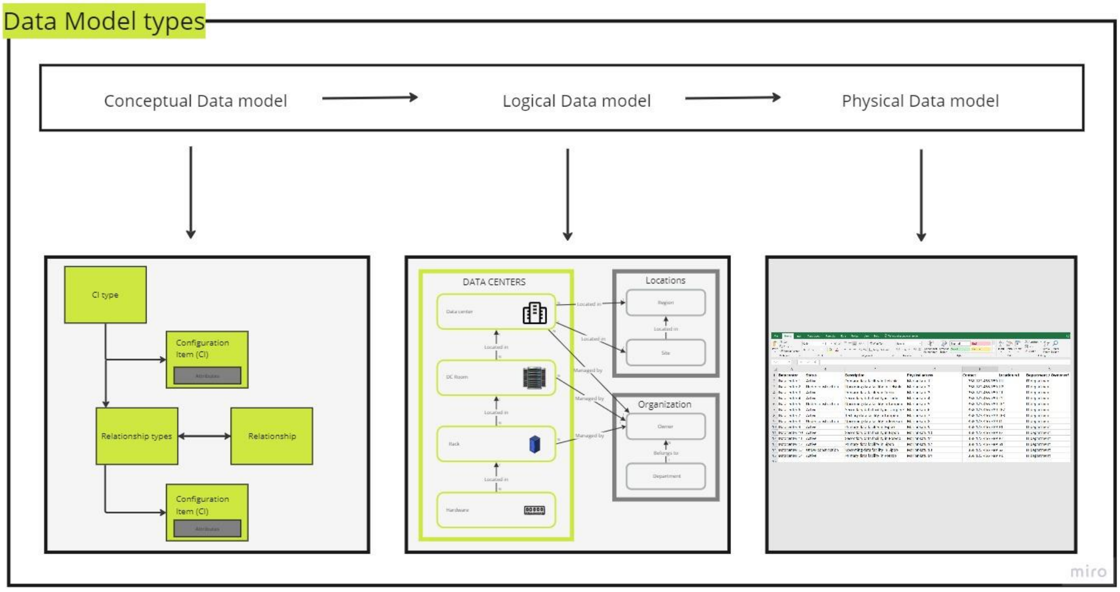

Figure 3, Data Model types. (Adapted from Justin Group Oy, 2020a & "Logical Design for Configuration Management Based on ITIL, " by H.Kyurkcheiv and K. Kaloyanova, 2012)

The complexity and design of a CDM are minimal. It captures the business entities, their high-level relationships, and their main attributes. The importance of the data model is to provide the bigger picture of the data and help align the work between business and IT stakeholders.

The LDM builds on upon the CDM and is a more detailed representation of the data model that will include data classes, fields, and relationships. These details will give the data model enough information to combine the technical and business views and a structured approach to represent the data logically.

The PDM describes and represents the implementation of the logical data model in a specific CMDB system. It is also crucial for the physical data model because it helps visualize the database structure by replicating the database column keys, columns, indexes, triggers, and other features (Justin Group, 2020b).

Figure 3 illustrates the progression from the CDM focusing on the high-level business entities, the LDM delving deeper into the relationships and dependencies, and the PDM showing how to arrange the data precisely within a CMDB system. This progression shows how every phase adds detail and complexity, ensuring the data model supports the company's CMDB.

2.2.4 Complexity in Data Models

The complexity of a data model in the CMDB refers to the degree of detail and intricacy of the CIs, the attributes, and the relationship between them. As stated, "They support the definition of templates (or class definitions) for CI records, which can contain types (int, char, boolean…) attributes. (Brenner & Gillmeister, 2014). The complexity comes from different factors, such as the number of CIs, attributes per CI, and the relationships between them. A complex data model will provide a more comprehensive system overview, making it suitable for tasks such as change management and risk assessment. However, it can also make the CMDB more challenging to use and update.

Therefore, it is essential to find a balance between complexity and usability. This balance means focusing on offering the most important parts, allowing CMDB designers to share and discuss ideas using common terminology, and limiting the complexity of the resulting CMDB (Brenner & Gillmeister, 2014).

While using visual tools to manage complexity by representing the data model in a digestible manner, techniques such as the Collective CI pattern, described by Brenner & Gillmeister (2014), significantly reduce complexity by using a single CI to represent many similar parts.

Moreover, there are many ways to set up a CMDB, as pointed out by Brenner and Gillmeister (2014), highlighting the immense need for an adaptable and standardized design to help with the CMDB design. This need aligns with the idea that using CMDB patterns can help the balance between utility and complexity in CMDB data models, leading to more effective and easier-to-understand CMDBs.

2.3 Bridging Theory to Practice

The consultant firm Justin emphasized the importance of a structured yet adaptable data model, aligning with practical insights rather than leaning strictly on theoretical frameworks. This practical approach goes hand in hand with the discussion on LinkedIn (2023), which elaborates on managing the complexity and scalability of CMDBs, diving into the importance of clearly defining the scope and objectives as well as the use of automation for an effective management process. Furthermore, the discussion touches on choosing an appropriate CMDB data model and structure, underlining the importance of a logical representation made by the company's needs.

Academic frameworks, such as those outlined in the ITIL, usually recommend a more standardized approach to the planning stage, focusing on the theoretical frameworks. Vendor recommendations like Device42 (2023) offer a balanced perspective, focusing on offering an accurate, consistent, and scalable CMDB. Justin's insights further explain the need for methods such as lectures, workshops, and conceptual frameworks. This comparison shows the need to combine academic and hands-on methods to make a solid CMDB planning process that meets the company's requirements.

3 Conceptual Framework

3.1 PDMC Abstraction Model

The Practical Data Model for CMDB (PDMC) is a template for creating data models inside the logical data model. The template is made by Justin Group (2020a), and it is a good starting point for making an adaptable data model. When implementing the template into an organization, it must be adopted and edited to fit in with the used database system and programs. The PDMC follows a layered abstraction model, as seen in Figure 4, designed to simplify the complexity of the IT environment by categorizing and grouping items into levels. It gives the reader a better understanding of the complex structures, even for non-technical stakeholders.

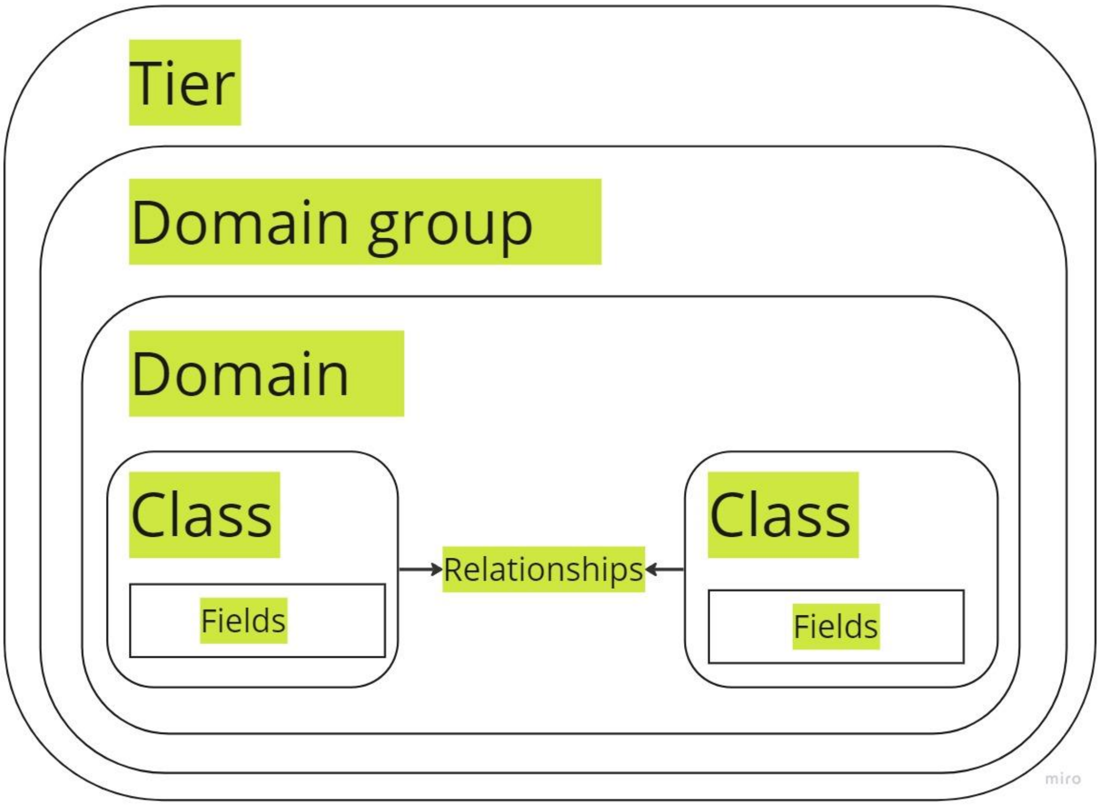

Figure 4, The PDMC abstraction model. (Adapted from Justin Group Oy, 2020b)

3.1.1 Tier

The tier stage, as represented in Figure 5 and detailed by Justin Group (2020a), splits into three stages: Business Tier, Service Tier, and Service Configuration Tier. The business tier represents the overall business goals and functions, the service tier includes both the customer-facing services and the backend services, and the service configuration tier delves into specifics of CIs and their relationships. Next to the primary tiers are two supporting tiers, which give essential data and oversee process management.

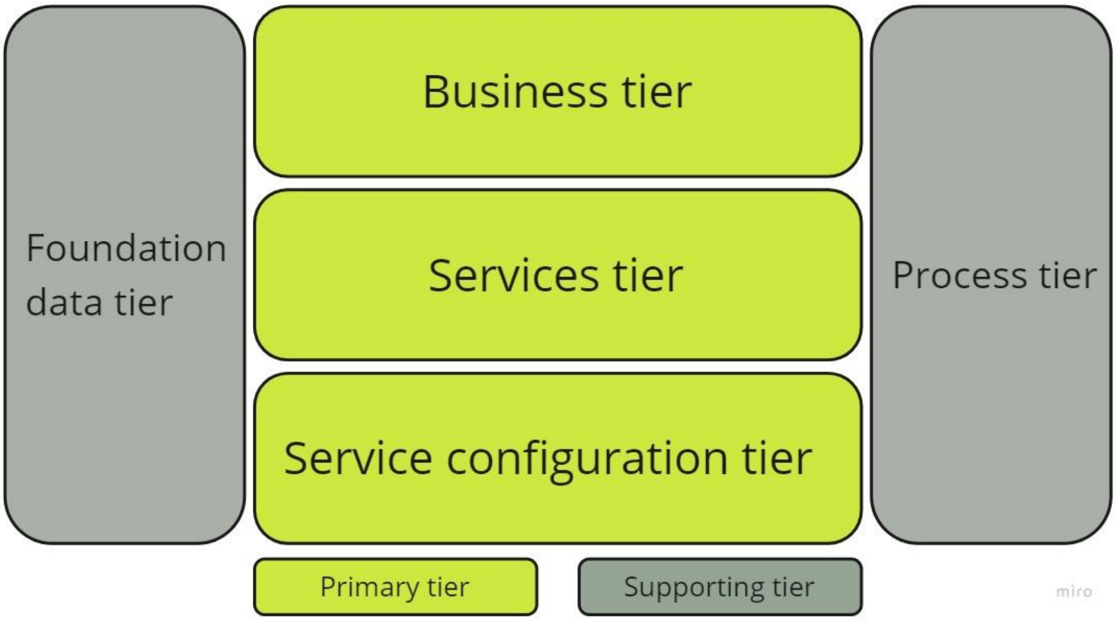

Figure 5, The PDMC abstraction model. (Adapted from Justin Group Oy, 2020a)

Business Tier

The business tier focuses on the highest-level services and processes that the company offers for its clients. It contains business services and business processes. These can include email services, car services, and insurance services. The second part is called the business process, a specific workflow to help the company achieve its goals, such as marketing, customer support, order processing, and sales.

Services Tier

The service tier combines the Customer Facing Service (CFS) and Backend Services. These services cover the part that the company is delivering to its customer and the backend part that supports the business operations. These CFSs can include online shopping platforms, customer portals, and consultancy services. The backend services include behind-the-scenes services such as payments gateway, inventory management, server hosting, and storage hosting.

Service Configuration Tier

The service configuration tier handles the detailed configuration of services and their components. It contains Frontend CIs and Backend CIs. The Frontend components form the user interface, such as software and interfaces. The backend components include servers, databases, and networking.

Foundation and process tiers

This tier provides a foundation of core data elements and supports management processes for the tiers. These are not actual parts of the CMDB or assets, but they are essential from an integrity and data management point-of-view. The foundation data tier is related to geography, organization, users, business, and finance. The process tier is incidents, changes, problems, and service requests.

3.1.2 Domain Group and Domains

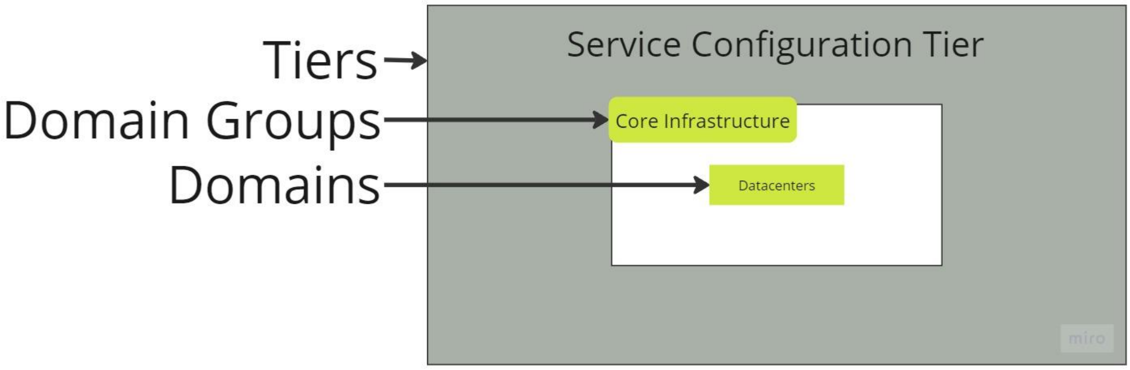

Figure 6, The PDMC abstraction model. (Adapted from Justin Group Oy, 2020b)

The domain group serves as a higher-level categorization or grouping technique. This group helps the organization to organize the different domains that they manage into categories. The data domains are a subset of the domain groups and are specific areas or categories within the IT environment. It can be seen as more personal than the domain group and focuses on specific types of CIs, data, or functions.

3.1.3 Class and Dependencies

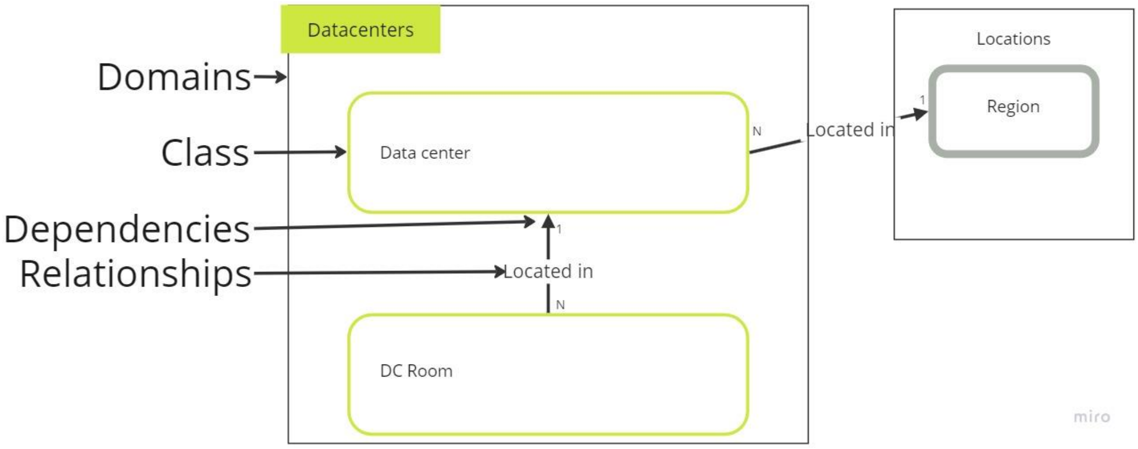

Figure 7, The PDMC abstraction model. (Adapted from Justin Group Oy, 2020b)

The data domains are detailed through diagrams with various classes and relationships of the specific data domain, as visualized in Figure 7. Classes are split into primary and supplementary classes, differentiated by colors. The primary class includes foundational attributes of, for example, the data centers. The supplementary class provides the primary class with additional information, such as the location. The coloring structure will help to keep the focus on the core domain while also recognizing the linked domains.

Class

The data classes are ways to categorize the data into similar types of things without worrying about where they are physically stored. For example, we have servers, network equipment, and storage devices. Even though these are in their categories or classes, they might share the same database table for efficiency reasons. This arrangement will make it possible to share the needed information based on the user. By planning the different categories in the logic data model, the data will run smoothly and help the personnel access the needed data.

Dependencies

Dependencies between entities are crucial to help ensure the database structure is efficient and prevent duplications. The types of dependencies, as outlined by The Support Group (2021), are:

- One to One (1:1): The entity is related to one record of the same or another data class and only one. For example, a server with a server identifier (like a Name or a Service Tag) will have a unique identifier, and each identifier will correspond to only one specific server.
- One to Many / Many to One, (*:1) or (1:n) as Justin (2020a) describes: Records from the primary data class can have many instances to other data classes. However, each instance of the other data classes is associated with only one instance of the leading data class. For example, a server rack can house multiple servers (One to Many). At the same time, each server is positioned in one specific rack (Many to One).
- Many to Many, (:) or (n:n) as Justin (2020a) describes: The record in one data class can have many instances with another data class or the other way around. For example, servers can host multiple different software applications, and in return, the software application might run on many different servers.

3.1.4 Fields

The fields are predefined slots for each data class where the users add the necessary data. Erillisverkot primarily uses the Netbox software and the fields from this system will be implemented in the CMDB. Netbox can define custom fields, allowing organizations to specify and capture data to their unique requirements. Erillisverkot will use the following typical field types during the implementation process (Netbox Documentation, 2023):

- Short text: a short description of the data (e.g., “Server Name” or “Asset Tag”).
- Extended text: a longer description or notes (e.g., “Servers specifications”).
- Dropdown list: predefined list of choices. (e.g., “Server type” with options like “Web” and “Database”).
- Date/Time: Capturing a specific date and time (e.g., “Installation Date” or “Warranty Expiry”).
- Number: Fields accepting only accept numbers (e.g., “Room number”).
- Toggle: Checkbox choice (e.g., “Active” with boxes Yes/No).
- Link/Reference: Connection to another data record (e.g., linking the room to a “data center”).

4 Methodology

4.1 Engagement Process with Justin

In a central role in the planning phase of the Configuration Management Database (CMDB) was the consultancy firm Justin. Their engagement spanned approximately three months, providing the key stakeholders, including myself, with essential lectures and workshops. Their process helped stakeholders understand the concept of CMDB and, more importantly, the planning of the data model, which laid a strong foundation for these research methods.

4.2 Qualitative Approach

Choosing the qualitative approach for this study was motivated by the amount of the requirements of the CMDB data model. The planning of a CMDB includes understanding complex organizational processes, infrastructure details, and the needs of the critical stakeholders, leading to the selection of a qualitative approach. “The qualitative method allows for creativity, varied interpretations, and flexibility. The scope of the research project can change as more information is gathered” (Grand Canyon University, 2023), essential from a CMDB data model point-of-view, where the model can evolve. Furthermore, the interviews, workshops, and meetings provided rich, detailed, and vital information for the data model.

4.3 Data Collection Methods

4.3.1 Observations

This study included observations of the interactive lectures and workshops conducted by Justin. These observations aimed to gain a better understanding of the CMDB concept and its application in the data model planning phase to align with the company’s needs.

Additionally, the observations also included meetings with experts in Data Center Infrastructure Management (DCIM). Although initially separate from the CMDB project, these meetings focused on agreeing on uniform naming conventions and the most important attributes, which go hand in hand with the attributes inside of the data model.

4.3.2 Interviews

The primary purpose of conducting the interviews was to better understand the core aspects and challenges of configuration management within Erillisverkot. A total of six stakeholders were interviewed via Skype, with the interviews lasting between 15 to 30 minutes, depending on the amount of the participant's ideas. Participants were chosen solely because of their involvement in the CMDB planning phase, and the objective was to get insight that could help shape the data model planning process.

The chosen method for the interviews was semi-structured. According to Barclay (2018) in KnowHow's learning resources for the KnowLife Partnership Hub, "a semi-structured interview is a qualitative research method that combines a pre-determined set of open questions (questions that prompt discussion) with the opportunity for the interviewer to explore particular themes or responses further". This approach made it possible to ask all necessary topics, with the questions constructed by Justin to understand the stakeholder's needs and expectations. The format gave the stakeholders more freedom to delve deeper into important topics and needs.

Among the different questions, the following key questions, translated from Finnish, were asked by head participants involved in the CMDB planning phase of the company from a data model perspective:

- *Why do we need configuration management?*

- *What information should the CMDB contain from my perspective, my team's perspective, and the entire company's viewpoint?*

- *Who would be the natural configuration data providers?*

- *Who would be the data consumers?*

- *Where is the configuration data currently maintained?*

The design of a possible data model comes from the conclusions of qualitative research, which includes interviews and observations, extensive literature review, and application of Justin's PDMC framework. The development methodology guarantees that the data model aligns with the requirements of the company and the challenges identified by stakeholders at Erillisverkot.

5 Results

5.1 Summarized Responses of the Interviews

- *Why do we need configuration management?*

Participants believed implementing a Configuration Management Database (CMDB) would give numerous advantages to help the company's operations. Based on the expectations of the interviewees, the data model should capture:

  - Device tracking: The participants explained a system that could comprehensively track core infrastructure components like servers, routers, switches, and firewalls, as well as personal equipment such as workstations, phones, and peripherals.
  - Enhanced visibility: This was a much-talked point in the interviews where the participants wanted a clearer view of the status, health, and performance of individual devices. One participant mentioned that such visibility could even allow for more significant error detection, fixing the issues before they affect the customer.
  - Support for processes: Participants mentioned that a CMDB would help with the different Information Technology Infrastructure Library (ITIL) processes, such as change management and incident and problem management. One interviewee noted, "It's crucial for our data center infrastructure that we can assess change management and incident management."
  - Mapping relationships: One mentioned problem within the company is that there needs to be an ability to map the relationship between

Configuration Items (CIs), with one participant stating, “Understanding how our server rooms interconnect is essential for risk management.”

- Lifecycle monitoring: The most mentioned thing by the interviewees was about lifecycle management, where it would be beneficial to keep track of the lifecycle stages of different CIs.
- Financial & technical insight: The person from the financial side mentioned the financial aspects of the CMDB, where it should be possible to get licensing costs and maintenance contracts.
- Other points: Other benefits mentioned during the interviews include greater quality and efficiency of services and less manual labor.

- *Who would be the natural configuration data providers?*

Natural data providers primarily include the service desk, team experts, and IT management. One participant mentioned, “The real-time changes need to be reflected straight to the CMDB.”

- *Who would be the data consumers?*

Everyone in the company will benefit from the data provided in the CMDB. Primarily, the teams that contribute the data will be its primary users. An IT manager stated, “From a planning point of view, the access to a centralized database is a game changer.”

- *Where is the configuration data currently maintained?*

The data is found across various platforms, repositories, and databases, such as Data Center Infrastructure Management (DCIM), Excel files, emails, and Jira. One of the participants noted, “We use a lot of time to find information from all the different sources.”

5.2 Observation Findings

5.2.1 Lectures

Key insights from the lectures to understanding the CMDB data model:

- Depth of Planning: The lectures showed that the planning phase of the data model is not just about a design but involves multiple layers. For a single Configuration

Item (CI) to be represented, multiple data models need to be constructed. For instance, for a single data center to be represented, there is a need for a higher data model to be present where all the data centers are located and have relationships between them.

- Scope: Properly defining the scope from the beginning helps the data model to be manageable and relevant, addressing the company's needs.
- Visual Aids: The use of visual representations in the lectures provided a clearer understanding of how data models should be made and how they interact.

5.2.2 Workshops

During the workshops held by Justin, participants engaged in essential discussions about translating theoretical knowledge into practical examples for the CMDB data model.

Key takeaways from these workshops:

- Depth of the Data Model: The hands-on workshops provide a comprehensive data model example that is not just a singular design but a layered construction. It was highlighted that representing even one CI was difficult and time-consuming, underscoring that multiple data models that are connected are needed.
- Company needs: These workshops were important for designing data classes and roles on the Mural platform to keep the data model aligned with the company's views. Participants were given the opportunity to engage in real-time interactions and provide feedback. One participant mentioned, "It would be ideal to construct the data model by prioritizing the components and excluding any unnecessary elements in the initial version."
- Data Model Structuring: The workshops, with the help of the stakeholders, gathered important information about the practicalities of the data model. The importance of a transparent, organized, and simple data model was highlighted, ensuring that everyone understands the information.
- Adaptable approach: The dynamic discussions during the workshops kept the data model flexible and ready for tweaks. In order for the company to progress it is essential to have an adaptable data model.

5.2.3 Meetings

The meetings made it clear to prioritize the development of the data center infrastructure data model to align with the company's preexisting DCIM system.

Key takeaways from these meetings:

- Refining the focus: In the meetings, narrowing down the data model focus to data center infrastructure was discussed, which aligned with the company's needs.
- Attribute identification: The most discussed matter during the meetings was pinpointing the most crucial attributes within the data centers and the hardware, ensuring that the data model would be relevant and efficient.
- Uniformity: Establishing uniform naming techniques was one of the topics, ensuring consistency across the data model.
- Alignment with the Needs: The meetings focused on aligning the data model planning phase with the needs of the company.

The participants were not directly part of the CMDB project, but they contributed insights during the meeting. These ideas were implemented in the planning and development of the data model.

5.3 Data Model Results

This section contains practical visual representations of the data models tailored for Erillisverkot data center infrastructure, offering a progression from a broad overview to detailed insights. Starting from the highest logical data model, highlighting the importance of data center infrastructure. This is followed by the logical data model, which delves into the key data classes and their relationships. Finally, the physical data model focuses on the main entities and attributes.

The first two models, the highest logical data model and the logical data model, are based on Justin's PDMC framework. These models have been extensively adapted and expanded to suit the requirements of the data center infrastructure at Erillisverkot. These adaptations are the result of theoretical concepts, frameworks, interviews, and observations, demonstrating the evolution of Justin's concepts to meet the needs of Erillisverkot.

The third model, the physical data model, is the key data model, focusing on the structure of the data center infrastructure within the CMDB.

5.3.1 Highest Logical Data Model

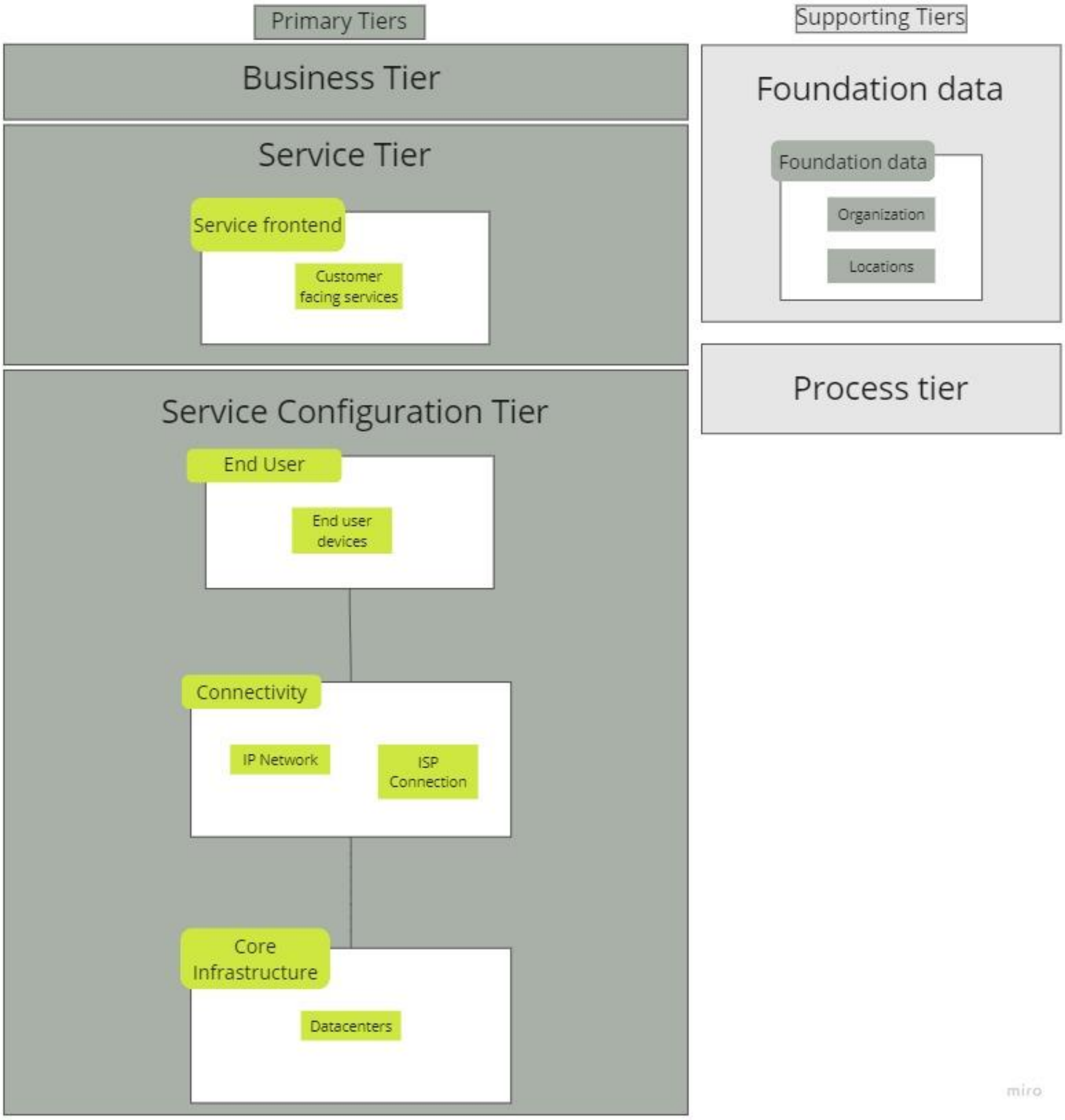

Figure 8, Conceptual representation of the highest logical data model for data center infrastructure. (Adapted from Justin Group Oy, 2020b)

Figure 8 lays out the conceptual representation outlining the plan for the company's data model. Considering the hierarchical approach, the diagram shows the progression from the overarching business tier down to the service configuration tier. The sections shown are the scope that the company will start with and then, with time, scale up and introduce more domain groups. The focused domain in this section will be the data center infrastructure, as it holds significant relevance to the operations of Erillisverkot.

5.3.2 Logical Data Model

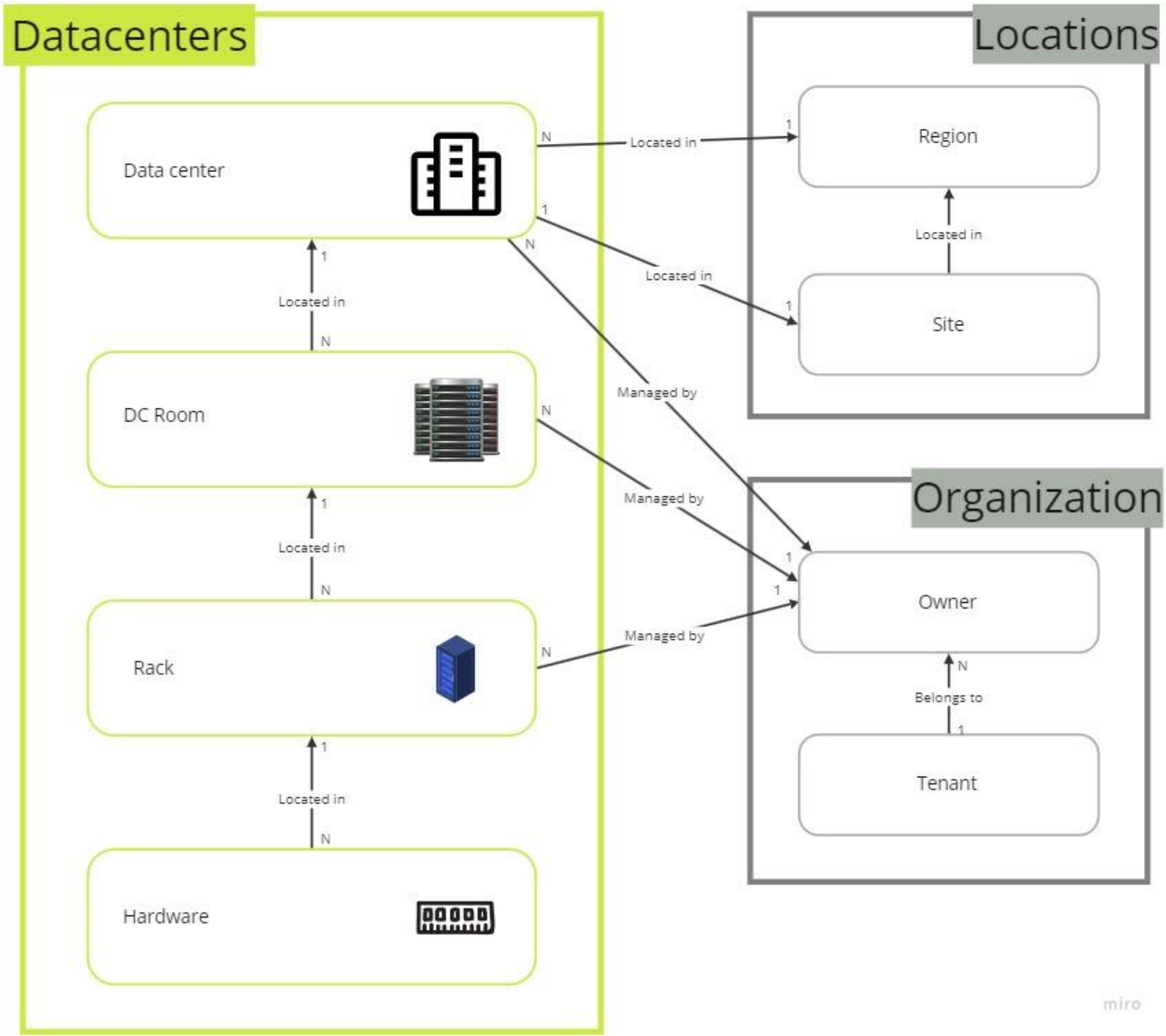

Figure 9, Conceptual representation of the logical data model for data center infrastructure. (Adapted from Justin Group Oy, 2020b)

Figure 9 delves deeper, providing a detailed view of the data center infrastructure blueprint. This Logical data model is specially tailored for the Erillisverkot Netbox structure. The key domains – data center, data center room (DC), Rack, and Hardware with the supporting tiers of Organization and Locations, come as important data classes for the company’s operations. Between these classes, there are the relationships and dependencies displayed, which help to understand the flow and to bind the system together.

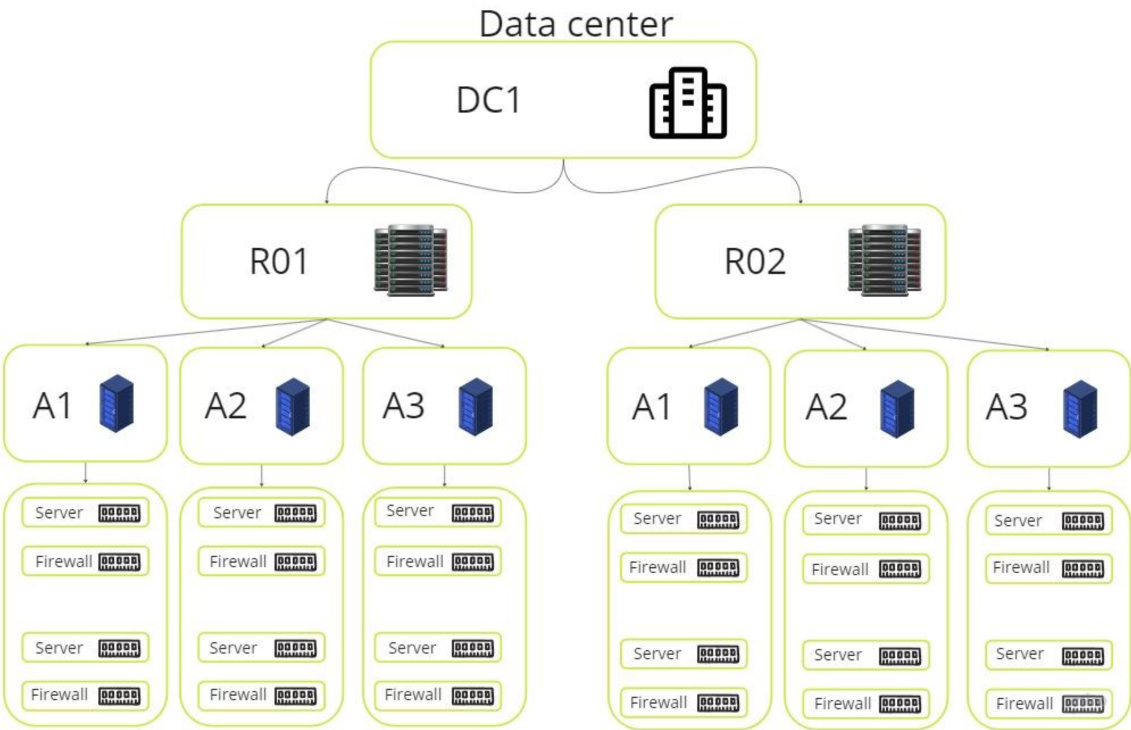

Figure 10, Data Center Infrastructure Hierarchy Diagram

Figure 10 is not a formal data model but a conceptual visualization intended to make a clearer understanding of the hierarchical structure within data centers. It maps out the arrangement between a singular data center connected to multiple rooms, racks within those rooms, and hardware devices inside each rack.

5.3.3 Physical Data Model

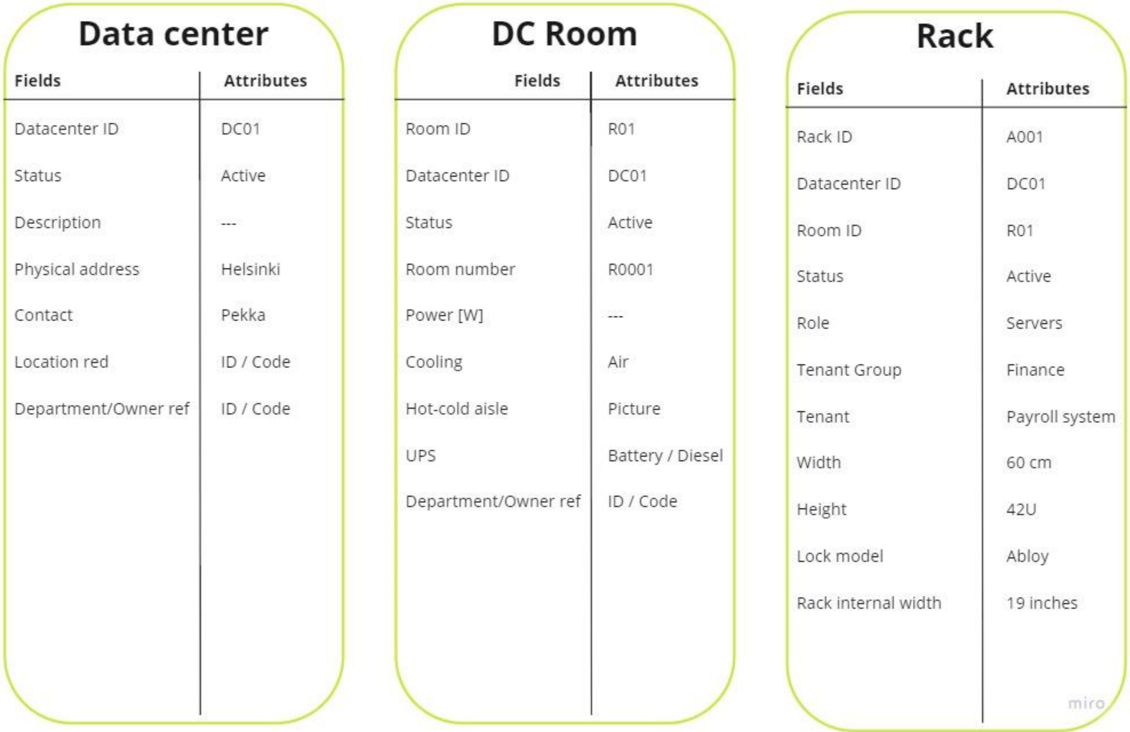

Figure 11, Conceptual representation of the physical data model for data center infrastructure: Detailing entity fields without the relationships

Figure 11, the physical data model, has been specifically designed for the Erillisverkot Netbox structure and will be used in the implementation of the CMDB. The model represents the main entities within the data center infrastructure – Data center, DC Room, and Rack. Each entity is detailed by its essential fields, with one example. While relationships are vital, this representation is focused solely on the entities and their attributes for clarity. The relationships between the groups would be the ID fields, which would be the point from where they connect.

6 Discussion

6.1 Evaluation of the CMDB Data Models for Erillisverkot

The evaluation of the CMDB data model for Erillisverkot reveals that it meets the company requirements in terms of the data center infrastructure. This model offers a detailed visualization starting from the highest logical data model and ending in the specific attributes of the data centers, rooms, and racks, the model provides a comprehensive overview of the infrastructure. The data model was heavily influenced by the knowledge gained from Justin's material, along with the semi-structured interviews and observations during meetings and workshops, ensuring that the data model was adaptable and actionable. This model ensures that stakeholders can easily understand and use the data for better decision-making.

Justin's Practical Data Model for CMDB (PDMC) framework was used for the development of the data models. The choice for Justin's framework was obvious because of their knowledge of CMDB's. This framework effectively demonstrates a structured yet straightforward approach by implementing knowledge of stakeholders and with the logic of Netbox system. As a result, the data model seamlessly integrates with systems like Netbox, serving as a blueprint for real-world applications.

The data model aligns with industry standards, such as ServiceNow's Common Service Data Model (CSDM), showing the adaptability to other platforms. This shows that the suggested data model could be used as a benchmark for other organizations. This comparison proves that the constructed data model and the framework are implementable in other systems, which is a key factor for scaling and adapting in an evolving IT sector.

The benefits of the developed data model:

- Fewer errors: The proposed data model simplifies the usage of CMDB by categorizing the Configuration Items (CIs) meaningfully. The reduced complexity makes the system more user-friendly and error-resistant.
- Bridges the IT-business gap: The proposed data model enables decision-makers to make better decisions from the data. It is a shared language between departments, allowing everyone to understand the model.
- Adaptability: The data model ensures that as software components evolve, the system can be adapted and expanded to new CIs, relationships, and attributes without significant changes.

6.2 Insights from Stakeholders Engagement

In a central role in the development process of the CMDB data model for Erillisverkot were the stakeholders. Through the series of lectures, semi-structured interviews, and meetings, rich knowledge and information were gathered, that significantly influenced the design of the data model. These discussions highlighted the need for a data model that provides comprehensive device tracking but also, in the beginning, an easy-to-understand model for every stakeholder.

From the interviews, it became clear that there was a strong need for a unified data model that would be able to effectively track devices to provide better visibility of the company structure and performance of equipment, while still supporting ITIL processes like change management. These insights added depth to the theoretical knowledge and showed the need for a data model that could fulfill these requirements.

The workshops often worked as collaborative brainstorming sessions for the stakeholders and played an important role in translating theoretical knowledge into practical. The sessions helped me understand the greatness of showing a single CI and the need to develop the data model in steps.

Regular meetings with stakeholders who were not directly involved in the CMDB project proved to be giving. These sessions helped align the data model with the existing system Netbox. The individuals during the meeting provided clarity on focusing the data model only on the data center infrastructure. They helped identify the most critical attributes for the data center, room, and rack while still keeping the model relevant and efficient. The importance of adding the locations and owner for the data center was one of the ideas that was discussed and implemented in the logical data model.

These methods underscore the importance of making the data model a step-by-step model, a key insight from Justin’s lectures. The approach was important when planning a possible solution for the data model that aligns with the company’s needs.

6.3 Challenges and Methodological Reflection

In evaluating the methodology, it is evident that the chosen approach was suited to the nature of this study. The methodology's adaptability and stakeholder-focused approach made it possible to navigate the challengers encountered.

One significant challenge was to find the most effective focus for the data model due to time constraints and the need to avoid over-complication. It became clear during the meetings that focusing on the data center infrastructure would provide a comprehensive yet manageable scope. This was important to offer a good data model while still keeping the planning process within feasible boundaries.

The main challenges were ensuring consistency and reliability of the received data acquired from the stakeholders during the interviews, workshops, and meetings. The participants had wildly varying levels of understanding about the CMDB, which made the responses both an opportunity for different insights and challenges for model unification.

Another challenge faced was the conduction of the interview in Finnish. Translating the answers and understanding the company-specific needs without losing their meaning required attention. Additionally, the help from the consultancy firm, Justin, was largely beneficial for the company's needs, but it did present its challenges. Aligning the needs for the company's data model with the consultancy material required communication and recalibration.

The examples of the potential data models for Erillisverkot were a blend of Justin's material, theoretical ideas, and practical insights. The challenges encountered during the process:

- Complexity of the model: The number of attributes, relationships, and CIs were a theme in this research. While the literature emphasized the challenges around complexity, the planning phase clarified the matter. Stakeholders discussed this in the workshops by addressing the challenge of representing even a single CI.

- DCIM attributes: During the planning phase, there was a dedicated effort to finding the best attributes within DCIM. The goal was clear: to craft a data model that uses the company's existing data and information. This topic became highly discussed during the meetings, where stakeholders had different opinions about which attribute was necessary for the data model.
- Insights from the stakeholders: The information gathered gave a glimpse into a potential data model structure, but it also came with complexities. The choice of the data center infrastructure as the data model example was brought to the table by one of the stakeholders.

6.4 Recommendations and Future Research

6.4.1 Recommendations for Implementation

While the planning process of the data model plays an essential role in CMDB, the implementation process will show things and problems that never reached the planning table. For this reason, the recommendations for the implementation process for Erillisverkot are:

1. Stakeholder Engagement: While in the planning and implementation process, regularly engage with the critical stakeholders of the project. They can provide deeper knowledge than any theoretical model can.
2. Continuous Lectures: After the CMDB implementation process, it is most important to give training to the personnel, especially those providing and using the data from the CMDB. The data model is designed to be able to adapt to different areas, and it is essential to keep the personnel up-to-date with the changes.
3. Integration with Existing Systems: Before choosing a platform for the CMDB implementation, it is important to choose software that supports the existing systems, like the DCIM system Netbox. This choice makes a smoother transition process.
4. Regular Reviews: Users must keep the data model updated and relevant. As the IT landscape and the company evolve, the model must be updated along with the changes.

5. Feedback: Inside the CMDB system, there should be a feedback mechanism where users can report issues, provide insights, and suggest improvements. Such a mechanism will keep the data model up-to-date and help problem management.

6.4.2 Suggestions for Future Research

The planning process for a potential CMDB data model for Erillisverkot shed light on the depth, challenges, and possible solutions. There are a few benefits that could be further explored:

1. Depth and User Friendliness: During the planning phase, the goal was to strike a balance between a detailed data model and one that is easy for users to navigate. It would be worthwhile to conduct research to delve deeper into finding the exact sweet spot between these two aspects in terms of real-world applications.
2. Stakeholders' feedback: Once the proposed data model is implemented, how will the stakeholders' ideas and needs change? It would be valuable to gather feedback after a period of real-world usage, as a follow-up as this could provide insights into needed upgrades.
3. Integration with new technologies: With new technologies emerging, how could the data model be adapted and improved by using AI and machine learning?

7 Conclusion

This thesis presents a plan that includes three detailed data models for the Configuration Management Database (CMDB) of Erillisverkot, a Finnish company specializing in secure ICT solutions. The research was driven by the challenges faced by Erillisverkot in managing the complex IT infrastructure and the need for a centralized database. This study aimed to plan a comprehensive CMDB data model of the data center infrastructure tailored to the company's specific needs.

The research was conducted using a qualitative approach, which included lectures, a series of semi-structured interviews, interactive workshops, and stakeholder meetings. This study reveals a complex combination of theoretical frameworks, company structure, and stakeholder perspectives.

The three proposed data models – the Highest Logical Data Model, the Logical Data Model, and the Physical Data Model – show a significant contribution to Erillisverkot IT infrastructure management. These models were designed to be adaptable, efficient, easy to understand, and in alignment with both theoretical frameworks and the company's practical needs, particularly the Netbox DCIM system.

In conclusion, this thesis gives more than just an academic understanding of the CMDB data model for Erillisverkot, but also gives actionable recommendations for implementation through the presented models. The study highlights the importance of engagement with stakeholders, training personnel in evolving environments, the need for compatibility with existing systems, and a need for regular reviews to maintain that the data model stays relevant and practical.

This study serves as an example for Erillisverkot or other organizations looking to enhance their CMDB data model. It lays the groundwork for developing CMDB data models even though these models are still conceptual and not yet implemented. They offer an approach for the planning process to address the challenges of IT infrastructure. Future application and testing of these models can further prove their effectiveness, serving as a potential guide for organizations seeking to enhance their IT infrastructure management and operational efficiency.

References

Barclay, C. (2018). *Semi-structured interviews*. KnowHow: Learning Resources for the KnowFife Partnership Hub. https://know.fife.scot/__data/assets/pdf_file/0028/177607/KnowHow-Semistructured-interviews.pdf

Brahmachary, A. (2018). *Service Knowledge Management System (SKMS) | ITIL Foundation | ITSM*. CertGuidance. https://www.certguidance.com/explaining-skms-itil/

Brenner, M., & Gillmeister, M. (2014). *Designing CMDB data models with good utility and limited complexity*. 2014 IEEE Network Operations and Management Symposium (NOMS). https://doi.org/10.1109/noms.2014.6838375

Device42. (2023). *The Technical Guide to CMDB Best Practices*. https://www.device42.com/cmdb-best-practices/

Drogseth, D., Sturm, R., & Twing, D. (2015). *CMDB Systems*. Morgan Kaufmann.

Grand Canyon University. (2023). *Qualitative vs. Quantitative Research: What's the Difference?* Doctoral Journey. https://www.gcu.edu/blog/doctoral-journey/qualitative-vs-quantitative-research-whats-difference

Groves, M. (2022). *What is Data Modeling? Conceptual, Physical, Logical*. Couchbase. https://www.couchbase.com/blog/conceptual-physical-logical-data-models/

Juola, M. (2023). *How to establish ownership in your CMDB*. Data Content Manager. https://datacontentmanager.com/how-to-establish-ownership-in-your-cmdb/

Justin Group Oy. (2020a). *JT06-PAC2.0* [Unpublished manuscript]. Received from Justin Group on May 30, 2023.

Justin Group Oy. (2020b). *The Practical Data Model for CMDB*. https://www.justin.fi/en/services/the-practical-data-model-for-cmdb/

Justin Group Oy. (2020c). *Services*. https://www.justin.fi/en/our-services/

Kyurkchiev, H., & Kaloyanova, K. (2012). *Logical Design for Configuration Management Based on ITIL*, Sofia University. https://www.researchgate.net/publication/251197977_Logical_Design_for_Configuration_Management_Based_on_ITIL.

LinkedIn. (2023). *How do you manage the complexity and scalability of your CMDB architecture and design?* https://www.linkedin.com/advice/3/how-do-you-manage-complexity-scalability-your-cmdb-architecture

Netbox Documentation. (2023). *Custom Fields*. https://docs.netbox.dev/en/stable/customization/custom-fields/

ServiceNow. (2023a). *CI relationships in the CMDB.* https://docs.servicenow.com/en-US/bundle/vancouver-servicenow-platform/page/product/configuration-management/concept/c_CIRelationships.html

ServiceNow. (2023b). *What is a common service data model (CSDM)?* https://www.servicenow.com/products/it-operations-management/what-is-csdm.html

Simplilearn. (2023). *Why CMDB - ITIL ®*. https://www.simplilearn.com/why-cmdb-itil-article#benefits_and_advantages_of_cmdb

The Support Group. (2021). *One-to-One and Many-to-Many Database Relationships.* https://blog.supportgroup.com/getting-started-with-relational-databases-one-to-one-and-many-to-many-relationships

8 Sammandrag på svenska

8.1 Bakgrund och syfte

I den digitala världen med komplexa IT-infrastrukturer är Configuration Management Database (CMDB) lösningen för en effektiv IT-tjänsthantering. Huvudsyftet med arbetet är att utforska och konstruera en potentiell datamodell för Erillisverkot kommande CMDB. En CMDB fungerar som en lagringsplats för företagets information och informationens relationer inom företaget, såsom hårdvara, mjukvara och nätverk som även kallas Configuration Items (CIs). Målet är att utveckla en strukturerad datamodellplan som effektivt hanterar informationen och relationerna.

8.1.1 Forskningsmål

Erillisverkot är ett statligt ägt finskt företag som erbjuder säkra ICT-lösningar för myndigheter och kritiska operatörer. De prioriterar säkerhet, pålitlighet och skräddarsydda tjänster. Forskningen kombinerar semistrukturerade intervjuer, möten och expertis av konsultföretaget Justin för att nå en potentiell datamodell för CMDB. Målet är att ge Erillisverkot en anpassningsbar och handlingsbar plan för datamodellen som fokuserar på datacenterinfrastrukturen. Studien fungerar som grundläggande steg mot implementeringen av CMDB-systemet hos Erillisverkot.

8.2 Forskningsmetodik

I planeringenfasen av CMDB datamodellen spelade konsultföretaget Justin en central roll. Under en period på 3 månader höll Justin föreläsningar och workshops för intressenterna i projektet, inklusive mig. Deras arbete underlättade förståelsen för CMDB-konceptet och var särskilt givande i planeringen av datamodellerna, vilket lade en stark grund för forskningsmetodiken.

8.2.1 Kvalitativ forskning och datainsamlingsmetoder

Studien använde sig av en kvalitativ metodik som anpassar sig till kraven av en CMDB datamodell. Studien kräver förståelse för komplexa organisatoriska processer, infrastrukturen och behoven av företaget. Den kvalitativa metodiken är flexibel vilket underlättade datamodellens uppbyggnad.

Datainsamlingen bestod av observationer av Justins interaktiva föreläsningar och workshopar, vilket ledde till en fördjupad förståelse av CMDB-konceptet och underlättad planeringsprocess. Forskningen inkluderade även möten med experter inom Data Center Infrastructure Management (DCIM) för att komma fram till enhetliga benämningar och identifiera de allra viktigaste attributen för enskilda CI inom datamodellen.

Genom semistrukturerade intervjuer samlades värdefulla insikter gällande utmaningarna inom konfigurationshanteringen hos Erillisverkot. Intervjuerna varade mellan 15 och 30 minuter var nyckeldeltagare, inklusive mig, inom CMDB-projektet deltog. Fokus under intervjuerna var att samla insikter som kunde forma planeringsprocessen för datamodellen, baserat på behoven och förväntningarna.

8.3 Resultat

Resultatet från studien visar flera fördelar med implementeringen av en CMDB hos Erillisverkot. Deltagarna underströk vikten av ett system för att effektivt uppfölja infrastrukturkomponenter som servar och routrar. Denna ökade synlighet hjälper till att övervaka enheternas status och prestanda, vilket identifierandes viktigt för risk- och förändringshantering samt incidenthantering.

Studiens observationer och workshoparna underströk behovet av en transparent, välstrukturerad och anpassningsbar datamodell som är lättförstådd. Tre huvudsakliga datamodeller presenterades:

1. Högsta Logiska Datamodellen (Figur 8): Utvisar en övergripande plan för företagets datamodell, inklusive olika sidonivåer.
2. Logiska Datamodellen (Figur 9): Ger en detaljerad syn av datacenterinfrastrukturen, med domäner som datacenter, serverrum, serverställ och hårdvara, samt relationerna mellan dessa.

3. Fysiska Datamodellen (Figur 11): Specifikt utformad enligt Erillisverkot Netbox-struktur, fokuserar sig på huvudattributen och fälten inom datacenterinfrastrukturen.

Resultatet av den kvalitativa forskningen, som inkluderade intervjuer, observationer, litteratur och en tillämpning av Justins PDMC-ramverk, säkerställde att datamodellerna stämde överens med företagets praktiska behov och klarade av de utmaningar som identifierades hos Erillisverkot.

8.4 Diskussion

Datamodellerna ger en detaljerad överblick av infrastrukturen som sträcker sig från den högsta logiska datamodellen, till de specifika attributen för ett enskilt datacenter som innehåller rum, serverställ och hårdvara. Utvecklingen av datamodellerna har varit starkt påverkad av Justins material samt av insikter från de semi-strukturerade intervjuerna och observationerna under möten och workshopen. Denna process har säkerställt att datamodellerna är i linje med företagets behov.

Utmaningar och reflektioner kring metodiken inkluderar hanteringen av komplexiteten i modellen och datainsamlingen av intressenterna. Översättningen av intervjuerna från finska och sammanställningen av Justins material krävde noggrannhet. Datamodellerna som presenterades kombinerar Justins teoretiska ramar med praktiska insikter. Resultaten visar betydelsen av att utveckla datamodellen stegvis och på basis av företagets behov.

8.5 Sammanfattning

Studien har utvecklat en plan som innehåller tre detaljerade datamodeller för Erillisverkot CMDB, vilket ger en effektiv förvaltning av IT-infrastrukturen. Den kvalitativa metodiken har kombinerat teoretiska och praktiska aspekter för att ta fram anpassningsbara och förståeliga modeller. Datamodellerna är en viktig del i hanteringen av komplexa IT-strukturer och erbjuder en strukturerad metod för att förstå och möjliggöra framtida utveckling inom IT-infrastrukturhantering.