Use of synthetic health data in prototyping for developing dental implant registry services

Oddmund Huseby

Master's thesis in Software Engineering at

Department of Computer science, Electrical engineering and Mathematical sciences,
Western Norway University of Applied Sciences

Department of Informatics,
University of Bergen

June 2022

Western Norway
University of
Applied Sciences

UNIVERSITAS BERGENSIS

Abstract

Developing novel applications in healthcare and dentistry can be challenging due to lack of application requirements, uncertain stakeholders, and no available test data. Such conditions exist in tooth implant dentistry, where innovative services are needed to record and communicate data. This study investigates how manually generated synthetic data can support the development and demonstration of new services for a dental implant registry. Furthermore, other objectives are to evaluate the usefulness of the developed services and to determine whether the development process has contributed to improving the data model used by a dental implant register.

To answer these objectives, we have through the use of design science methodology developed a high-fidelity dashboard prototype and a synthetic dataset in parallel. The development process was carried out in iterations, involving stakeholders as early as possible. The results indicate that the use of synthetic data to demonstrate possible future services was an essential component of the development process, facilitating early active participation of stakeholders. In particular, data with some realistic qualities were the most valuable in this process. Furthermore, the development process we used resulted in some contributions to the registry's data model, but fewer than expected. In summary, the services we developed were deemed useful by stakeholders.

These results suggest that synthetic data generated manually, together with a high-fidelity prototype, may contribute to involving stakeholders early in the development process. This participation may ease the process of identifying application requirements and engaging stakeholders, potentially producing useful features.

Acknowledgements

First and foremost, I would like to thank my supervisors, Yngve Lamo and Svein-Ivar Lillehaug, for their invaluable advice and guidance. Thanks should also go to the other members of the project team, Trine Lise Lundekvam Berge, Stein Atle Lie and Gunvor Bengtun Lygre from the Department of Clinical Dentistry at the University of Bergen and Norce Research, who generously provided from their knowledge and expertise. I am also grateful to all the people who helped me evaluate my work. Lastly, I would like to thank my family for their encouragement and support.

Contents

List of Figures

List of Tables

Acronyms

CDSS Clinical Decision Support System.

EHR Electronic Health Record.

HIT Health Information Technology.

MVC Model-view-controller.

ORSS Outcome Assessment and Research Support System.

PROM Patient-reported outcome measures.

SUS System Usability Scale.

UI User Interface.

Chapter 1

Introduction

In this chapter, we first present the significance and value of the investigation. Subsequently, we describe the research problem in more detail before listing our research questions and introducing the methods we use to answer them. Then we clarify related work and describe the contribution of the project. Finally, we outline the content and structure of the thesis.

1.1 Motivation

Health Information Technology (HIT) is recognised as a crucial factor in improving the quality of care, and its use has increased in the last decade (Charles et al., 2015; Mahajan et al., 2021). The use of HIT may transform the delivery of care by increasing safety and effectiveness, positively affecting medical outcomes (Kruse and Beane, 2018; Shekelle et al., 2006; Chaudhry et al., 2006). A form of HIT that has gained popularity is clinical and quality dashboards (Dowding et al., 2015). There are indications that this technology positively impacts patient care if implemented appropriately (Dowding et al., 2015; Khairat et al., 2018; Oliver et al., 2019).

However, developing dashboards within healthcare is a highly complex issue that requires careful consideration of handling, integration, and presentation of data (Ghazisaeidi et al., 2015). Unfortunately, few projects use evaluation as part of the design process. Instead, they perform it after the deployment to validate the performance (Zhuang et al., 2020). In general, this issue of delaying user involvement and evaluations appears to be recurrent in HIT (Yen and S. Bakken, 2012). Consequently, some negative impacts usually associated with reduced user involvement can affect these projects, such as increased costs, undiscovered design flaws, and less relevant and usable end-products (Alvertis et al., 2016).

Consequently, to ensure that HIT, such as dashboards, meet the needs of their users, stakeholder participation must occur at all stages (Hartzler et al., 2015). This participation also includes the early stages of development. In this phase, there may be uncertain application requirements (Marinho et al., 2014) and reluctant stakeholders unsure of the new possibilities of the HIT (e.g., Meskó et al. (2017)). Engaging stakeholders and obtaining feedback from them as early as possible can mitigate some of these issues (Maqbool and Herold, 2021). In this process, it might be beneficial to discuss more than conceptual ideas and let users interact with a functional prototype. Involving and immersing stakeholders is probably important, especially when the artefact emphasises data exploration (Mannino and Abouzied, 2019; Marieke McCloskey, 2014). It may be even more critical when novel applications are developed within a new field. In many cases, however, this approach requires a functional prototype populated with adequate test data. Unfortunately, there is a lack of such data in healthcare due to privacy and legal concerns (Maqbool and Herold, 2021). Similar fields handling patient data may face the same problem. Consequently, it is challenging for developers to perform demonstrations and evaluations as a part of the development process within such domains.

An option to acquire the necessary development and testing data is to use synthetic data (James et al., 2021). Synthetic data implies "data that are artificially created/simulated" rather than being generated by actual events" (Dilmegani, 2022). Mainly, there are two strategies to generate such data. First, create data that mimic the properties of a real data set, for example, through Generative Adversarial Networks (Goodfellow et al., 2014). Second, develop data without any existing data set, using publicly available information and statistics to guide the process or define the distributions manually. A patient record generator that uses this approach using public statistics is *Synthea* (Walonoski et al., 2018).

In developing novel services, actual data might not exist, leaving only the option of generating synthetic data without any existing data set. In these cases, there is a chance that there are no patient-record generators or the generators are insufficient. Consequently, the developers' only option to obtain data is to generate them manually. This creation of variables and their distributions and relationships can be complex, making the acquisition of test data challenging. This challenge may cause developers to bypass early feedback and reduce the number of demonstrations and evaluations.

Consequently, it will be valuable to gain an understanding of whether manually generating a synthetic data set, in parallel with the development of a dashboard prototype, can assist in the development process by engaging stakeholders. This assistance may facilitate the identification of possible functionalities, services, and requirements in collaboration with stakeholders. To test this approach, we iteratively develop the synthetic data set and some possible dashboard services, highlighting the challenges and solutions we discover along the way. The purpose of the synthetic data set is to explore how it can contribute to and influence software development, i.e. the development of new services. The overall goal of the dashboard is to (1) demonstrate potential new services for the future use of quality registers by presenting possible functionality, and (2) engage stakeholders by providing insights into new opportunities.

1.2 Problem Description

Implant dentistry is an example of a field with limited available data and opportunities for innovation. Several studies have proposed a systematic collection of relevant data on dental implants (e.g., a quality registry) to improve the quality of treatment and identify implants that perform better (Klinge, Lundström, et al., 2018; Naemi et al., 2021; Pye et al., 2009). Medical/health registers contain individualised data related to the patient's treatment process, and their purpose is to continuously improve the quality of care (*Quality registries* 2021). Registries achieve this goal by providing researchers with aggregated data collected at the regional and national levels in various areas of healthcare. In addition, selected statistics from this data are summarised and presented in annual reports, which highlight areas of possible improvement for hospitals and clinicians, providing them with actionable information to improve patient care (e.g., Jernberg et al. (2014), Ludvigsson et al. (2019), Fredriksson et al. (2017), and *Quality registries* (2021)). There are indications that dashboards can further enhance the usefulness of quality registers (Dowding et al., 2015). Dashboard services may provide faster information communication, interactive visualisations, and analysis capabilities.

Developing and creating novel applications, such as dashboard services, within a new field can be challenging. Currently, there are no dental implant registries, although, the title of a recent systematic review, "Dental implant quality registries and databases: A systematic review" (Naemi et al., 2021), might give the impression that such registries exist. A study of this paper does, however, uncover that the title is misleading as none of the included studies in this review report on any quality registries. Furthermore, to our knowledge, there are no dashboards or other services that communicate data from other types of dental registries. Although there have been some attempts to assess the quality of dental care using simple queries against electronic records (Neumann et al., 2017), there are no such examples for dental implants. This context creates a challenging development environment, predominantly due to three factors:

- The requirements for a dashboard are uncertain because it is the first of its kind.
- Stakeholders are unaware of new possible user services.
- Data are not available for demonstration and evaluation.

As demonstrations and a continuous evaluation process during the development process may assist in uncovering an application's requirements and engaging stakeholders, the three factors will influence each other. Therefore, acquiring a data set to use in the process of developing a prototype may be vital to creating a useful dashboard. As we did not discover any suitable data generation tool, we must manually define the variables and their distributions and relationships. Such a manual approach can be daunting for data-intensive applications that handle many variables. As a potential solution, we propose an iterative development process, creating two artefacts in parallel: the dashboard and the synthetic data set. This solution can potentially simplify the task of producing the synthetic data, as it is possible to adapt it to our needs throughout development.

In combination, these artefacts act as a high-fidelity prototype that can showcase potential functionality to stakeholders. We reckon that this approach is ideal for developing data-intensive applications without having access to real data since the data necessary for computations and visualisations are created during development. Additionally, there may be other benefits to using this approach. It may reveal new requirements for the system and the underlying data model because it is possible to explore the data through the dashboard. In addition, it can engage users who are unsure of the possibilities related to novel registry services.

An important aspect to consider during this process is the qualities that can be incorporated into the synthetic data. For example, a higher degree of realism in the data may influence its usefulness in the development process.

The intriguing possibilities of a dental implant registry combined with a supplementary dashboard and the challenges of developing such an application provide the ideal context for testing our approach. Our research group just finished creating a data model for a dental implant registry (Vågenes, 2022). Consequently, this project's natural starting point is to use this model to create a synthetic data set, which acts as a foundation for building the dashboard.

RQ1 How can synthetic data generated manually support the development and the demonstration of new services for a potential dental implant registry?

RQ2 What is a possible useful implementation of secondary services for a dental implant registry?

RQ3 How can the process of generating synthetic data manually and creating secondary services contribute to the improvement of the data model used by a dental implant quality registry?

1.3 Research Methods

In this project, our objective was to develop a prototype and create a data set to perform demonstrations and evaluations. Hence, it was natural to use the design science paradigm to develop our artefacts.

1.3.1 Design science

At the heart of design science lies a rigorous process of designing and evaluating artefacts to solve specific problems (Dresch et al., 2015). As design science is a paradigm, it does not propose a detailed step-by-step process for conducting design science research (Hevner and Chatterjee, 2010). However, an essential aspect of a design science project is understanding the research cycles.

The *relevance cycle* connects the environment with design science research. Often, the process begins by identifying opportunities and problems that provide the requirements and acceptance criteria for the study (Hevner, 2007). Hevner (2007) emphasise that researchers must return the artefacts produced by design science research to the environment for field testing and evaluation. This reintroduction will help determine whether the initial assumptions are correct and whether further iterations are necessary.

The *rigour cycle* links the design science research with the knowledge base, allowing researchers to draw from existing ideas (Hevner, 2007). According to Hevner (2007), a vital aspect of this cycle is to ensure that the artefacts produced are not a routine design but contain innovations. This assumption is

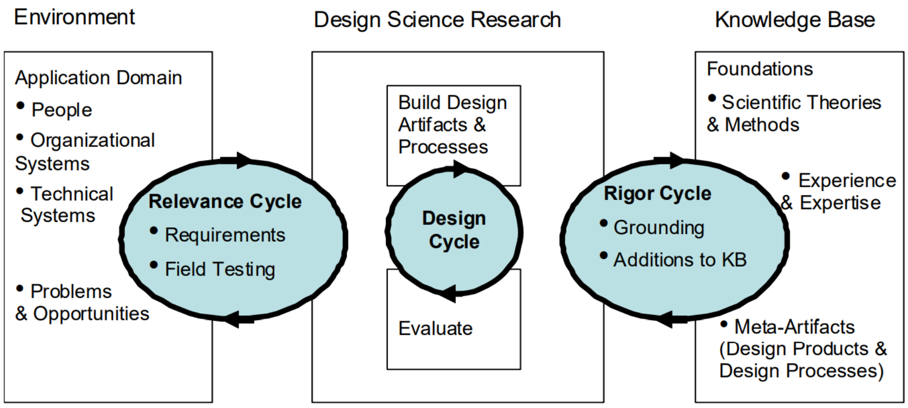


Figure 1.1: The research cycles that connect the different elements. From Hevner (2007)

critical because design science should contribute to the knowledge base. This contribution may be new additions and insight into original theories, unique artefacts, and new experiences gained throughout the design process (Hevner, 2007).

The *design cycle* occurs within design science research, rapidly generating between the construction of the artefact, its evaluation, and its refinement (Hevner, 2007). An example is to develop different alternatives and evaluate them against the identified requirements. This internal validation occurs before an artefact is released to the *relevance* or *rigour cycle* and ensures that it maintains a certain quality (Hevner, 2007).

1.3.2 Design science research

The three cycles must be visible and identifiable in a design science research project. Peffers et al. (2007) presented a methodology to conduct design science research to help researchers. The process includes six steps:

1. Problem identification and motivation
2. Define the objectives for a solution
3. Design and development
4. Demonstration
5. Evaluation
6. Communication

We use these activities to guide and evaluate the design research project. The first item, *Problem identification and motivation*, is part of the *relevance cycle* and defines a research problem and justifies the solution's value (Peffers et al., 2007). This thesis presents the motivation in Section 1.1 while listing the research problems in Section 1.2.

The following section presents step 2, the *objectives of the solution*. The *development process*, which includes steps 3, 4 and 5, we describe in more detail in Chapter 3. The last step, *communication*, we perform throughout the thesis.

Define the objectives of the solution

The research questions do not necessarily translate directly into goals for the artefacts. Concequently, it may be helpful to specify the objectives of the solution, for example, by describing how we expect the artefact to support solutions to unaddressed problems (Peffers et al., 2007). Hence, we list three objectives:

1. **Synthetic data** - Determine an approach to generate a synthetic data set that can be used in a functional prototype, supporting the development of registry services when real data are unavailable.

2. **Registry services** - Identify potential features and functionality that may be helpful to stakeholders in exploring data from a dental implant registry.

3. **Synthetic data and registry services** - Evaluate how the parallel development of the services and synthetic data may contribute to the registry's data model, further developing and verifying it.

We will develop registry services using a dashboard. These services may benefit various stakeholders by providing easy access to data related to dental implants. Dashboards are well suited for the task of communicating data to stakeholders due to their interactive visualisations and analysis capabilities. By finding solutions to the first two objectives, future registry services can benefit from easier development.

Interestingly, parallel development of the dashboard and the synthetic data set may improve the registry beyond the two artefacts. The solution to objective 3 may influence the data model and other parts of the system. Exposing stakeholders to the synthetic data set through the dashboard, thereby simulating the registry's data content after years of usage, may help them better understand possible secondary services and data requirements. This information may help uncover potential problems or identify new opportunities.

We decided to develop a feature-rich dashboard prototype for a dental implant registry and populate the registry with synthetic data to solve these objectives.

1.4 Related Work

There exist no study directly related to the work we performed in this project. However, several studies are partially related to our research.

- Pollack et al. (2019) explore the possibility of using synthetic data created manually to test various dashboard designs. Their goal was to identify the best design to solve a particular task. Domain experts assisted in the creation of the data, deciding all the values of their data set. For comparison, we generated data with less detail on specific values. In addition, we used the data earlier and for a different purpose; engaging stakeholders and identifying potential requirements for future applications.

- Nelson et al. (2016) developed a dashboard to facilitate health information exchange between two departments. Similarly to our approach, they used prototyping to engage stakeholders and identify their application requirements. However, they used a combination of low- and high-fidelity prototypes. Moreover, their high-fidelity prototype was not functional. We present more of this study in Section 2.6.
- In healthcare, several dashboards have been created (e.g., Stattin et al. (2016), Dagliati et al. (2018), Lindblad et al. (2016), and Weiss et al. (2018)). Although the domain is different, we may use some of the lessons in our project. These studies are presented in more detail in Section 2.5.

1.5 Contributions

This thesis makes several contributions to the literature. First, it defines a new way of using manually generated synthetic data in development, showing how developers can employ it with a high-fidelity prototype to develop new services for a dental implant registry. Second, the result of this project helps to provide a better understanding of the qualities of synthetic data that contribute to the development process. Finally, it adds to the many dashboard examples seen in the literature. Since it is the only one for dental implants, its functionality may be an inspiration for future implementations within this field.

1.6 Outline

The outline of the thesis is as follows:

- **Introduction** - Chapter 1 presents the motivation and describes the problem. We also introduce design science, which we use to investigate our research questions. Consequently, we translate our research questions into objectives for the solutions we develop.
- **Background** - Chapter 2 contains information on synthetic data and its use in software development. Additionally, we introduce implant dentistry, quality registries, and dashboards, and we discuss how quality registries and dashboards may contribute to implant dentistry. Finally, we present prototyping as a potential approach for developing registry services.
- **Methodology** - Chapter 3 describes the design process we used to develop the synthetic data set and the dashboard.
- **Design** - Chapter 4 presents each iteration and its activities, describing the evolution of the two artefacts. We used the same structure for each iteration, starting with defining tasks and finishing with evaluating the implementations.
- **Implementation** - Chapter 5 presents a description of the artefacts that includes the tools we used to develop the artefact, features, architecture, and other relevant attributes.
- **Final evaluation** - Chapter 6 presents a comprehensive review of the artefacts. First, we summarise the findings of the semi-structured inter-

views, which assess how the artefacts fulfil our objectives. Second, we identify possible stakeholders of the different features and review the synthetic data's realism and the application's performance.

- **Discussion** - In Chapter 7, we discuss our findings.
- **Conclusion** - Finally, in Chapter 8, we summarise the key points of our thesis.

Chapter 2

Background

In this chapter, we will briefly introduce synthetic data generation, addressing the challenge of developing novel applications that lack actual data. Subsequently, we describe implant dentistry, highlighting the need for structured patient data within this field. Additionally, quality registers, combined with a visual analytic dashboard, are presented as a potential solution to some of the challenges in implant dentistry. We inform about the current use of these technologies, along with relevant definitions and case studies. Finally, we present prototyping as a potential development approach for creating dashboards.

2.1 Synthetic Data

There is a growing interest in synthetic data, with use cases ranging from machine learning to internal software testing (Zerdick, 2021; James et al., 2021). This technology provides benefits such as the reduced risk of privacy infringements and financial and technological advantages (James et al., 2021). Synthetic data can be defined as "data that are artificially created/simulated" (Dilmegani, 2022). There are several ways to produce synthetic data, and we can separate the strategies into two groups. Some generation strategies require an existing data set, replacing the original values but capturing the complexities and relationships in real data (e.g., Reiter et al. (2014) and Goodfellow et al. (2014)). On the other hand, some strategies do not require actual data but may use publicly available information and statistics to guide the process or define the distributions manually (e.g., Walonoski et al. (2018)).

Furthermore, the term synthetic data may imply *fully/purely* or *partially/semi*-synthetic data. There is no exact definition, but *fully/purely* synthetic data may indicate that the entire data set is generated (Hahner et al., 2019). In comparison, *partial/semi*-synthetic data combine the generated and real data in the data set (Hahner et al., 2019). Therefore, depending on the use case and the requirements, stakeholders can employ different generation strategies and combinations. For example, when training AI models for diagnosis and prognosis, the synthetic data should closely resemble the observational data from the real world (Chen et al., 2021).

Synthetic data are well suited as a tool in software development, enabling testing and demonstration without risking privacy violations. Depending on the application and its current stage in production, developers' needs regarding the degree of similarity between synthetic and actual data may vary (James et al., 2021). For example, a mixture of random strings and numeric values can be sufficient for early development (Mannino and Abouzied, 2019). Later, developers will likely need relatable and informative data to demonstrate the application to stakeholders (Mannino and Abouzied, 2019). For complex systems where data drive much of the behaviour, the data must be statistically representative and logically valid (Soltana et al., 2017). Alternatively, if the goal is to evaluate an application's specific features and design choices, the data must be applicable for scenario testing (Pollack et al., 2019).

Hence, synthetic data can aid in various aspects of software development. A challenge occurs when developers have to create synthetic data themselves because of the lack of actual data or data generators. It is important to note that by generating synthetic data manually, we mean that a creator sets the parameters and other attributes of the data, and a programme generates it according to these settings. Such circumstances may arise when developing innovative applications within a new field. The difficulty in manually implementing the statistical properties of synthetic data is to ensure a proper relationship between the variables and introduce errors and missing values (Mannino and Abouzied, 2019). Although there are tools that can support the creation of realistic-looking data (e.g., *Synner* and *simstudy* (Mannino and Abouzied, 2019; Goldfeld and Wujciak-Jens, 2020)), manually generating data can be a daunting task. Consequently, developers might settle for simplistic and fragmented test data during development. Such data may be helpful to some extent but may reduce stakeholder involvement and engagement.

Domains, such as dental implants, currently lack actual data and patient-record generators. Hence, developing new services and applications may be challenging due to inadequate test data. In the following sections, we will introduce implant dentistry and dashboards, highlighting how the use of new technology can improve dental implants and procedures. Afterwards, we present prototyping as an approach that developers may use to create novel applications. This approach, combined with a synthetic data set, can potentially assist in developing services for a dental implant registry.

2.2 Implant dentistry

Dental implants have gained widespread adoption and are the preferred treatment for missing teeth (Buser et al., 2017). The use of dental implants has increased in recent decades (Jenny et al., 2016). Globally, more than 100 different brands sell approximately 12-18 million implants per year (Klinge, Lundström, et al., 2018). The unique characteristics of implants and other external factors can influence their success rate (Lee et al., 2005; Raikar et al., 2017).

The overall success rate of dental implants is high, and several studies have observed very few or no complications after 5 years (Cochran et al., 2007; Beschnidt et al., 2018). However, researchers have no consensus on what constitutes the best implant characteristics, such as the shape and roughness of the implant surface (Elias, 2011). The attributes of the patient, the characteristics of the implant, and the implant's position all affect the survival rate of dental implants (Raikar et al., 2017).

Consequently, dentists must consider numerous factors when selecting an appropriate dental implant for a specific patient. For example, the position of the new implant will probably affect the implant size chosen by the dentist (Lemos et al., 2016). According to Lemos et al. (2016), short implants are likely to be selected as a treatment for posterior jaws. This selection can reduce the complexity of the insertion but can present a higher risk of failure later (Lemos et al., 2016).

An emphasis on evidence-based practise is needed to navigate this environment. However, there is often a lack of scientific evidence before new implants and procedures are implemented, and few studies examine the long-term outcome (Klinge, Lundström, et al., 2018). According to Finkelstein et al. (2020), the lack of evidence-based decisions in dentistry can be partially attributed to the limited availability of data. Even so, this field is increasingly collecting and integrating more data, opening new possibilities within dental research (Kusiak and Somerman, 2016; Joda et al., 2018). Hence, stakeholders must record all relevant data to more rigorously evaluate procedures and implants, for example, using a quality registry (Klinge, Lundström, et al., 2018; Pye et al., 2009).

2.3 Quality registries

A quality registry contains individualised data related to the patient's treatment process, including diagnosis, treatment outcomes, and patient-reported outcomes. Each registry covers members of a subpopulation with a particular condition. In general, the main purpose of collecting all these data is to continuously improve the quality of care (Adami and Hernán, 2015; *Kvalitetsforbedring* n.d.). To achieve this goal and maintain its status as a quality registry, it must satisfy specific requirements. Below is a list of some of the criteria that a certified medical quality registry at the highest level in Norway must meet (*Stadieinndeling* n.d.):

- The validity and reliability of the collected data must be documented.
- At least 80% coverage of patients eligible to be recorded.
- Collect Patient-reported outcome measures (PROM).
- Data must be used in research.
- Annually publish quality indicators.
- Document that registry analyses have been used to identify areas that need improvement.
- Document that the registry's results are used in quality improvements, either by the registry itself or by cooperating with clinicians.
- Publish results from the registry and adapt them to different target audiences.

Based on this list, a quality registry develops the quality of care indirectly or directly through research or using the registry's results in quality improvements. Below is a description of how these two approaches may contribute to the quality of care.

- **Research** - Research using data from quality registries can uncover new information that stakeholders can use in quality improvements. The completeness and validity of the data and its longitudinal properties enable studies that are not possible with randomised clinical trials (Furu et al., 2010; Kieler, 2010). For example, the duration of drug exposure can differ substantially in a short investigation compared to actual clinical practise (Furu et al., 2010). Therefore, using data from a quality registry can reveal how that drug functions in the real world (Furu et al., 2010). Journals have published more than 200 articles based on data collected from Norwegian quality registries (I. J. Bakken et al., 2020).
- **The registry's result** - One method quality registries use to communicate their results is annual reports. These reports are published to relevant stakeholders and positively affect the quality of care. Hartmann-Johnsen et al. (2019) discovered that hospitals used such documents to improve practise, resulting in increased usage of recommended treatments. Jernberg et al. (2014) did the same observation, with around 90% of the hospitals reporting using annual reports from a national quality registry to identify areas that need improvement. Furthermore, other studies have found similar results, suggesting that annual reports communicating the results of the registry can help develop the quality of care (Ludvigsson et al., 2019; Fredriksson et al., 2017).

Hence, a dental implant quality registry could potentially improve the quality of care by providing actionable information to stakeholders. The benefits of a quality registry for dental implants are recognised by Klinge, Sanz, et al. (2018), which lists the perceived advantages and disadvantages. The disadvantages are data protection, fear of being compared and scrutinised, and a higher administrative burden. Relevant stakeholders must consider these aspects thoroughly. On the other hand, perceived advantages range from identifying good practices and detecting problems with certain implants to more patient involvement.

Some of the main advantages of a dental implant quality registry, or a registry in any domain, are the rapid assessment of new treatments and new technology and the examination of the performance of different dental implants (Klinge, Sanz, et al., 2018). Therefore, appropriate tools are needed to provide fast and reliable information to the various stakeholders. Annual reports used in health care are valuable but introduce some delays from recording to using the data. Santos and Eriksson (2014) highlighted this challenge and indicated that potentially significant variations in data could go unnoticed and that important outliers may be overlooked when data are published infrequently. Furthermore, the static nature of reports may cause significant trends to go undiscovered (Ghazisaeidi et al., 2015). A possible solution is to publish reports more frequently (Santos and Eriksson, 2014). However, a more feasible solution to reduce the overhead of generating reports may be to use a dashboard. The real-time capabilities of this tool and the additional possibilities of visual exploration and manipulation of data may complement annual reports. Hence, the dashboard may assist in communicating the registry's result and thereby more fully utilising the potential of a quality registry.

2.4 Dashboards

It is challenging to define a dashboard accurately, as researchers apply it to many different entities (Sarikaya et al., 2019). Few (2006) highlight that one of the critical characteristics of a dashboard is to communicate important information quickly to the user. Other definitions have also included characteristics that encompass the user's purpose and understanding (e.g. Yigitbasioglu and Velcu (2012)). In this project, we will use the following definition as given by Zhuang et al. (2020):

> "A dashboard is a visual information system which comprises at least a graphical user interface (GUI) and a store of data which the GUI exposes, which is designed and built with the purpose of fulfilling a precise information need, and whose primary information transfer channel is a responsive display."

This definition describes some of the characteristics of a dashboard and calls attention to the dashboard's purpose. Thus, creating a dashboard is more than just creating a GUI; it is necessary to define a goal. In some cases, the intent may be evident. However, the final objective may be unknown when developing completely new services. This uncertainty requires iterating the purpose during development to find the niche of the new service and its related requirements.

2.4.1 Types of dashboards

Part of the definition involves fulfilling a precise information need. Developers must identify the purpose and adapt the dashboard to its requirements, which is critical for a successful implementation (Ghazisaeidi et al., 2015; Clarke et al., 2016). Typically, a dashboard has a set of tasks that guide users toward satisfying their information needs. Zhuang et al. (2020) define these as "intended tasks". The determinability of the outcome and the expected number of user interactions influence the complexity of a task (Zhuang et al., 2020). By using this categorisation, Zhuang et al. (2020) divide the dashboards into four different types:

- Decision support: The tasks have high determinability but may require a lot of interaction by the user. The interactions could be setting other parameters and thresholds, which would result in a fixed outcome. Various practitioners use this type of dashboard to aid their decision process, for example, by highlighting important information to the user.
- Operational: This type of dashboard requires minimal interaction and is highly deterministic. The purpose may be to inform users about the status of specific key performance indicators. The dashboard has low complexity but has the advantage of delivering near real-time updates.
- Exploratory: As its name indicates, this dashboard gives a range of possibilities to the user. Typically, the user can explore the data, searching for patterns and unusual observations. Therefore, it may require a lot

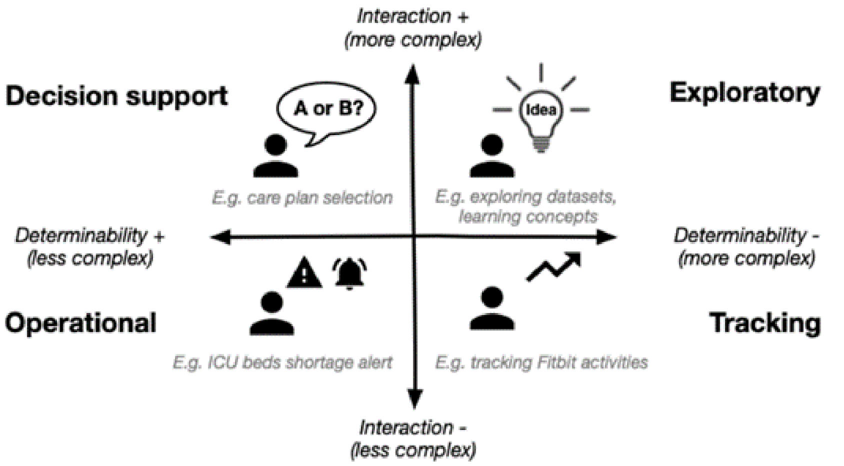
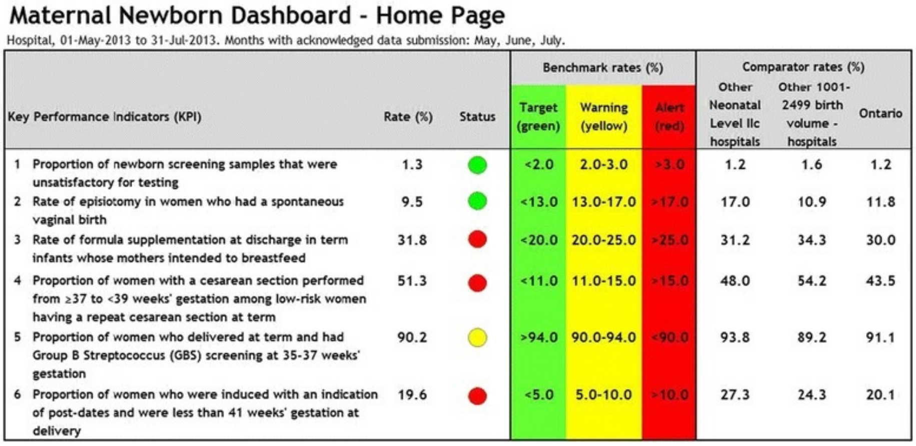


Figure 2.1: Four different dashboard types. The two axes display the complexity of various operations. Interaction is related to the number of user interactions, while determinability is related to the liberties given to a user exploring the data. Illustration from Zhuang et al. (2020).

Figure 2.2: A clinical decision dashboard that summarises the performance of treatment alternatives, giving the clinicians options to specify the importance of specific attributes (Dolan et al., 2013).



Maternal Newborn Dashboard - Home Page

Hospital, 01-May-2013 to 31-Jul-2013. Months with acknowledged data submission: May, June, July.

	Key Performance Indicators (KPI)	Rate (%)	Status	Benchmark rates (%)			Comparator rates (%)		
				Target (green)	Warning (yellow)	Alert (red)	Other Neonatal Level IIc hospitals	Other 1001-2499 birth volume - hospitals	Ontario
1	Proportion of newborn screening samples that were unsatisfactory for testing	1.3		<2.0	2.0-3.0	>3.0	1.2	1.6	1.2
2	Rate of episiotomy in women who had a spontaneous vaginal birth	9.5		<13.0	13.0-17.0	>17.0	17.0	10.9	11.8
3	Rate of formula supplementation at discharge in term infants whose mothers intended to breastfeed	31.8		<20.0	20.0-25.0	>25.0	31.2	34.3	30.0
4	Proportion of women with a cesarean section performed from ≥37 to <39 weeks' gestation among low-risk women having a repeat cesarean section at term	51.3		<11.0	11.0-15.0	>15.0	48.0	54.2	43.5
5	Proportion of women who delivered at term and had Group B Streptococcus (GBS) screening at 35-37 weeks' gestation	90.2		>94.0	90.0-94.0	<90.0	93.8	89.2	91.1
6	Proportion of women who were induced with an indication of post-dates and were less than 41 weeks' gestation at delivery	19.6		<5.0	5.0-10.0	>10.0	27.3	24.3	20.1

Figure 2.3: An operational dashboard showing key performance indicators for maternal-newborn care (Dunn et al., 2016). In addition, the dashboard shows benchmarks to provide direction for practice change.

of interaction and may present open-ended outcomes. Successful use is time-consuming and involves training and detailed domain knowledge.

- A tracking dashboard tracks different metrics over time. The user often has the slight possibility for any interaction but may exert some cognitive effort to try and interpret the data.

2.4.2 Evaluation of dashboards

In recent years, the popularity of dashboards within healthcare has increased (Randell et al., 2019). To have a rewarding experience interacting with the dashboard, users rely on it to perform its intended task adequately. Therefore, many studies advocate for a user-centred design approach and ensure adequate user training (Wright et al., 2019; Franklin et al., 2017; West et al., 2014; Randell et al., 2019).

However, projects often do not evaluate dashboard design during the development process but instead perform post-deployment evaluations (Zhuang et al., 2020). Zhuang et al. (2020) advocate using multiple evaluation criteria to continuously assess the dashboard during the design process to ensure good feedback, including interaction effectiveness, user experience, and system efficacy. In many cases, testing these aspects requires a usable application with data. Consequently, creating a prototype to involve users as early as possible may be beneficial.

2.5 Quality registries and dashboards

One of the purposes of quality registries is to use their data to develop the quality of care. On the other hand, dashboards communicate data to fulfil a precise information need, utilising visualisations and interactivity. Together, dashboards can improve or supplement data reporting and findings from quality

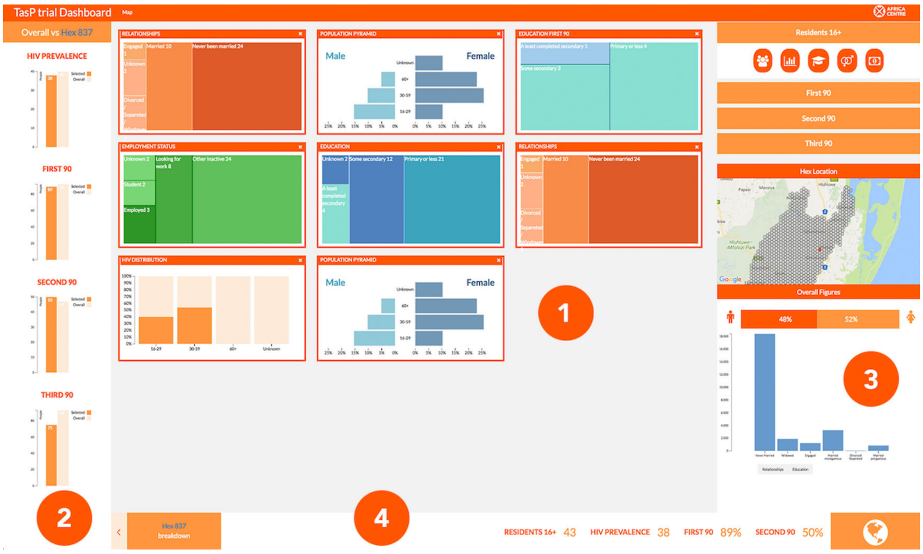


Figure 2.4: An analysis page of an exploratory dashboard visualising data sets produced at a demographic health surveillance site (Concannon et al., 2019). Users can add or remove visualisations and extract key performance indicators.

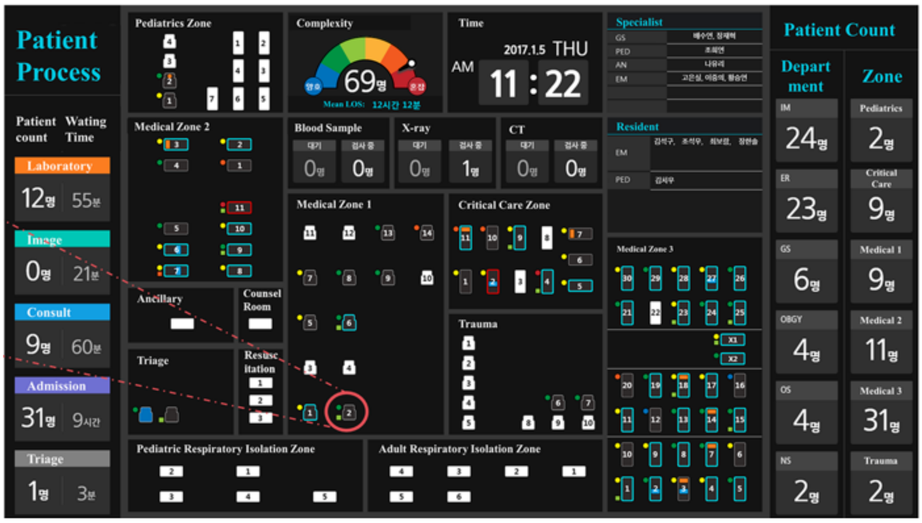


Figure 2.5: A tracking dashboard showing statistics regarding an emergency department's performance in real-time (Yoo et al., 2018).

registries. This usage may be beneficial to clinicians and scientists using the information produced by the registry.

Currently, there are no quality registries in implant dentistry, and to our knowledge, there are no visual analytic dashboards to communicate information within dental care either. Consequently, this might cause a reduction in the flow of information to relevant stakeholders, as well as delaying potential valuable research and quality improvements.

Meanwhile, dashboards have already been used in clinical and quality improvement settings in health care, reporting mainly data from Electronic Health Record (EHR) or various registries (Dowding et al., 2015). Although the impact of some studies is less known, most studies report positive results (Weiss et al., 2018; Dowding et al., 2015). According to Dowding et al. (2015), dashboards may improve adherence to quality guidelines, improving patient care. Factors such as dashboard design and ease of access appear to be influential in achieving these results (Dowding et al., 2015).

Some of the knowledge and insights of studies that combine dashboards and EHR or registries may be useful and adaptable to dental implant dentistry.

2.5.1 Dashboards reporting quality indicators

In healthcare, dashboards have been used to report quality indicators based on data from quality registries. One of the goals of this approach is to present data as precise and actionable information to healthcare providers (Stattin et al., 2016). The near real-time capabilities of a dashboard can help reflect the current quality of care and reduce reporting delays, which could make it a valuable tool for improving care (Stattin et al., 2016). Weiss et al. (2018) investigated the effect of one such dashboard (see Figure 2.6). The dashboard was part of a registry that collected pregnancy, birth, and childhood data in Ontario, Canada (Reszel et al., 2019). The dashboard reported six key performance indicators to increase practitioners' awareness concerning the hospitals' quality of care (Reszel et al., 2019). They selected the indicators due to their importance in patient outcomes (Dunn et al., 2016). The dashboard was of the tracking type, with little interaction but some cognitive effort to interpret the result. The dashboard evaluation compared the quality indicators before and after introducing this tool. The result was statistically significant improvements for four of six key performance indicators. Therefore, the dashboard was associated with better care for mothers and babies.

We can adapt this idea of highlighting key indicators for various stakeholders in this project. Since Stattin et al. (2016) used this approach to compare hospitals, we could apply the same method to dental clinics. Such a dashboard would likely draw attention to essential aspects of the quality of care in the different clinics, urging practitioners to improve and continuously informing them.

2.5.2 Exploratory dashboards

There are examples of the deployment of more advanced dashboards within healthcare. These types of dashboards can be used in clinical decision support

	Key Performance Indicators	Rate (%)	Status	Benchmark rates (%) Target (green)	Warning (yellow)	Alert (red)	Comparator rates (%) Other Neonatal Level IIIa/IIIb hospitals	Other >4000 birth volume hospitals	Ontario
1	Proportion of newborn screening samples that were unsatisfactory for testing	2.6		<1.0	≥1.0 and <1.5	≥1.5	0.6	1.5	1.4
2	Rate of episiotomy in women who had a spontaneous vaginal birth	13.5		<13.0	13.0-17.0	>17.0	5.9	9.3	8.2
3	Rate of formula supplementation from birth to discharge in term infants whose mothers intended to exclusively breastfeed	1.6		<20.0	20.0-25.0	>25.0	16.4	25.3	23.8
4	Proportion of women with a cesarean section performed from ≥37 to <39 weeks' gestation among low-risk women having a repeat cesarean section at term	4.2		<11.0	11.0-15.0	>15.0	19.0	41.6	36.9
5	Proportion of women who delivered at term and had Group B Streptococcus (GBS) screening at 35-37 weeks' gestation	53.4		>94.0	90.0-94.0	<90.0	98.1	92.8	91.1
6	Proportion of women who were induced with an indication of post-dates and were less than 41 weeks' gestation at delivery	67.1		<5.0	5.0-10.0	>10.0	9.5	6.6	17.3

Data source BORN Ontario, 2016-2017

Figure 2.6: Sample of the Maternal Newborn Dashboard from Reszel et al. (2019)

and to manage the patient treatment process. Dagliati et al. (2018) developed a system with two types of dashboards to control the treatment process of type-two diabetes. They directed the Clinical Decision Support System (CDSS) dashboard to the care providers. The goal was to assist them in exploring the risk of developing complications in patients. The Outcome Assessment and Research Support System (ORSS) dashboard was directed at healthcare managers and policymakers to support decision-making and research (Figure 2.7). Dagliati et al. (2018) highlight that the successful aspects of these dashboards were that they combined data from multiple sources in a single view. The CDSS dashboard led to improved metrics such as reduced visitation duration. This improvement was due to clinicians having a better understanding of the patient's condition. Health care managers found the ORSS dashboard helpful in exploring patient data and deviations from routines.

Different views and exploratory features are highly relevant to our project. This kind of dashboard would allow stakeholders, such as researchers and practitioners, to explore the data, possibly helping them identify valuable trends in dental implants and procedures.

2.5.3 Dashboards prompting cooperation

While dashboards can display quality indicators or allow one to explore data sets, their potential use is even more comprehensive. Recently, healthcare has moved from focussing on disease treatment to increasing attention to patients' preferences, concerns, and goals (Batalden et al., 2016). In this process, a visual analytical dashboard can facilitate collaboration and co-production between caregivers and patients, which was the focus of a study by Lindblad et al. (2016). Specifically, they used a real-time dashboard that used patient-reported

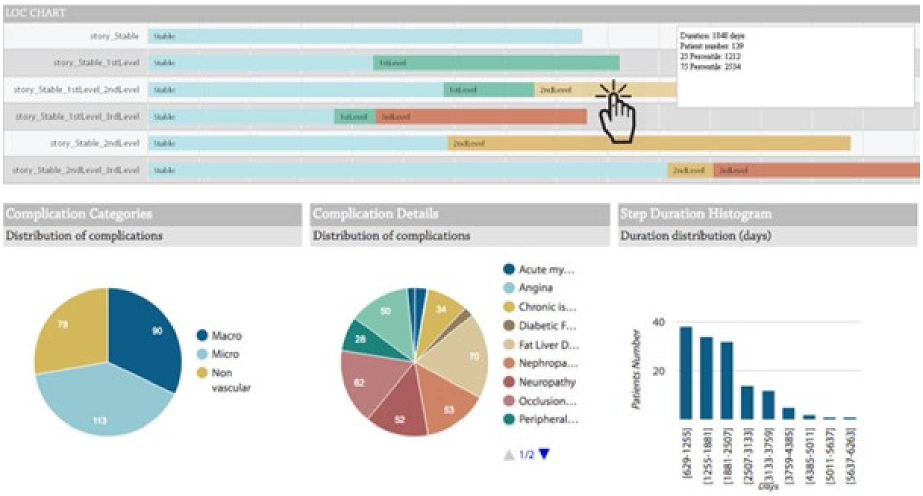


Figure 2.7: Sample of the outcome assessment and research support system dashboard from Dagliati et al. (2018)

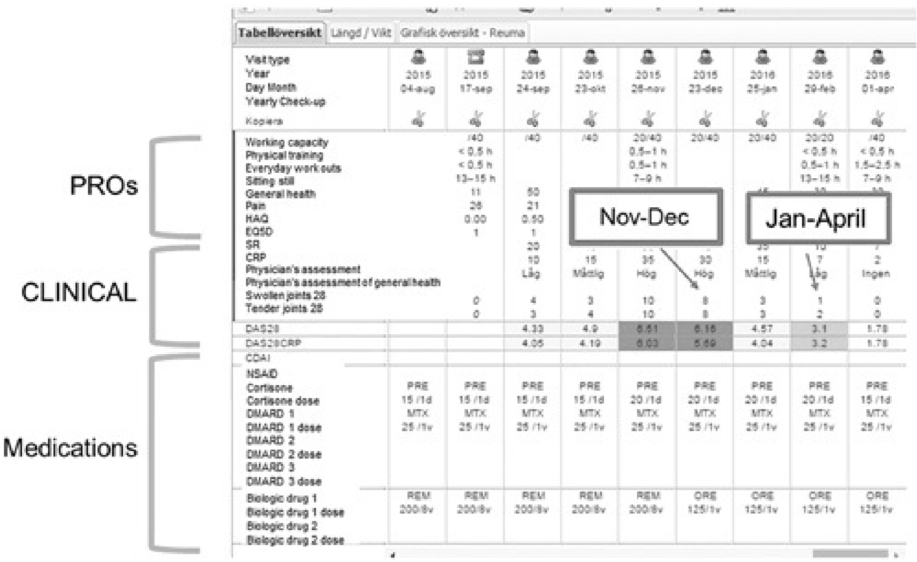


Figure 2.8: Show an example of the co-production module which caregivers and patients view to facilitate cooperation and shared decision making. The display integrates patient-reported outcomes (PRO), clinical data, and treatment history from the quality registry. From Oliver et al. (2019)

data along with clinical information to promote joint decisions (Lindblad et al., 2016). This dashboard integrated data from the Swedish Rheumatology Quality Registry and had three components (see Figure 2.8).

First, it had a patient-facing component that tracked patient-reported outcomes between visits. Therefore, patients could monitor aspects of their health and infer how treatments and other actions affected it (Lindblad et al., 2016). The clinical team used the second component to view patient-reported data, outcomes, and treatment trends (Lindblad et al., 2016). The third and last component was the co-production component, which many clinicians believed to be the most valuable feature of the dashboard. It facilitated the conversation between patients and clinicians by showing patient-reported results and clinical data, possibly comparing subgroups (Lindblad et al., 2016). Using these tools, patients and clinicians could develop a treatment plan and evaluate it at each encounter based on the observed results and trends (Lindblad et al., 2016).

Consequently, the treatment plan was continuously monitored and tailored to the patient. A study by Oliver et al. (2019) summarised some of the findings of this registry, reporting a reduced level of rheumatoid arthritis in the population. They contributed this change to several factors but credited the various dashboard components with some improvements.

2.6 Prototyping

Prototyping is an activity that creates a preliminary version of an application to gain feedback on the software requirements (Bourque et al., 2014). Typically, the prototype implements a selected subset of the properties of a system that helps to acquire relevant input from stakeholders (Kordon and Luqi, 2002). Such simplifications facilitate a faster and easier build process, which is essential for an effective prototype (Kordon and Luqi, 2002). When developing a prototype, developers must consider the prototyping style, the target, and the evaluation technique (Bourque et al., 2014):

- **Prototyping style** is related to the various approaches to creating prototypes. According to Davis (1992), there are mainly two ways to develop prototypes:
  - The throwaway approach is used to quickly prototype the least understood parts of an application. The code is discarded after it has completed its mission.
  - The evolutionary approach constantly refines a prototype, producing a robust prototype that becomes part of the final product. One of its use-cases is to uncover application requirements.

  However, there exist other methods as well and various combinations. For example, in a study by Escalona et al. (2021), they attempt to reuse more of the knowledge gained from throwaway prototypes.

- **Prototyping target** refers to the target that benefits from the activity. This target can be the requirements specification, architectural design, or other aspects of the application.

- **Evaluation Techniques** specify how developers or other relevant stakeholders evaluate the prototype. The evaluation may test the prototype against an actual application or a list of requirements.

Prototyping can be used in different development approaches, for example, *Rapid Prototyping.* This approach is iterative and aims to rapidly improve the design and functionality of applications. According to Merrill (2018), *Rapid Prototyping* consists of three steps. Developers first create a high- or low-fidelity prototype, then evaluate it with relevant stakeholders, and thirdly use the feedback to improve the prototype and start again at step two.

The advantage of this method is that it allows early user feedback, which can help to discover application requirements (Merrill, 2018). These requirements can be related to usability, functionality, and other aspects. Gordon and Bieman (1995) found that rapid prototyping improved ease of use and helped identify user needs. Therefore, this might be an excellent method to develop novel services with unclear requirements.

In healthcare, there are examples of projects that have used rapid prototyping to develop dashboards. De Croon et al. (2015) used this method to create a proof-of-concept dashboard, which contributed to integrating user feedback into the application. Similarly, rapid prototyping ensured that useful feedback from stakeholders was incorporated into a project by Nelson et al. (2016). In this project, which sought to refine the requirements for a health information exchange dashboard, they registered these additional benefits of prototyping:

- Motivates users to contribute ideas for improvements
- Domain experts involved with the design
- It was easy to identify shortcomings
- Showcasing possible designs promotes innovation

Consequently, there seems to be ample opportunity to use prototyping in the process of developing secondary services for a dental implant quality registry.

2.6.1 Section summary

The use of synthetic data has increased, providing benefits in various fields. In software development, synthetic data may provide access to test data when actual data are restricted due to privacy or no available data. Therefore, developers can access the data needed to build and test applications. An area with limited access to actual data is implant dentistry.

The low amount of data in implant dentistry may partly explain the lack of evidence-based practise with respect to dental implants and procedures. By recording all relevant data, a more rigorous assessment is possible. Quality registries have been suggested as a tool to record data. In health care, these registries allow research and help identify areas of improvement. They usually communicate results through annual reports that hospitals, managers, and clinicians use diligently. Dashboards show promise in supplementing annual reports by providing real-time capabilities and possibilities for visual exploration and manipulation of data.

Quality registries and dashboards can quickly report data and findings, allowing stakeholders to assess and survey different dental implants' performance rapidly. Therefore, such technology can innovate the field of implant dentistry. Developing a dashboard for a quality registry can be a significant challenge, especially due to the lack of available test data for demonstration and evaluation, the lack of detailed application requirements, and stakeholders not knowing of the possibilities.

To try and counter these challenges, we used an approach in which we created a synthetic data set and a dashboard prototype in parallel, allowing us to use a high-fidelity prototype early in the development process. This method can contribute to early stakeholder participation, uncover application requirements, and assist in discovering adjustments to the underlying data model.

Chapter 3

Methodology

In this project, part of our goal was to create an innovative dashboard to demonstrate, motivate, and highlight new possibilities of secondary services of a dental implant registry. To identify the requirements of such an application, we wanted to use prototyping, which may facilitate early user participation and engagement. The prototype was high-fidelity since much of the application's functionality would rely heavily on data from the registry. However, the challenge was to develop this prototype in a new field where no actual data was available. Therefore, we explored an approach to creating two artefacts in parallel using the design science paradigm: (1) a synthetic data set populating a dental implant registry and (2) a dashboard to communicate the data stored in the registry.

We wanted to study whether this parallel development benefited each artefact by facilitating greater user participation and making it easier to identify valuable features and issues, resulting in a more practical application for the dental implant registry. In addition, we wanted to examine how the realism of the synthetic data influenced this approach, for example, whether more data with more realistic qualities were more helpful in determining the system's needs.

To understand how the artefacts are connected, we will first present the system's high-level architecture. Afterwards, we present the approach and methods that we used to develop the artefacts.

3.1 High-level architecture

Figure 3.1 shows the data flow between the system's various components, indicating the connection between our two artefacts. In general, we built the system using the FLUX pattern developed by Facebook, which employs a unidirectional flow of data compared to Model-view-controller (MVC) (Tay, 2019). At the start of a user session, the server requests all data from the dental implant registry and then processes and stores them in the application store. Next, the store notifies the view, and the view queries the necessary data and displays them to the user. The grey lines in Figure 3.1 show the chain of events that occur when a user interacts with the application. A dispatcher, which is a kind of

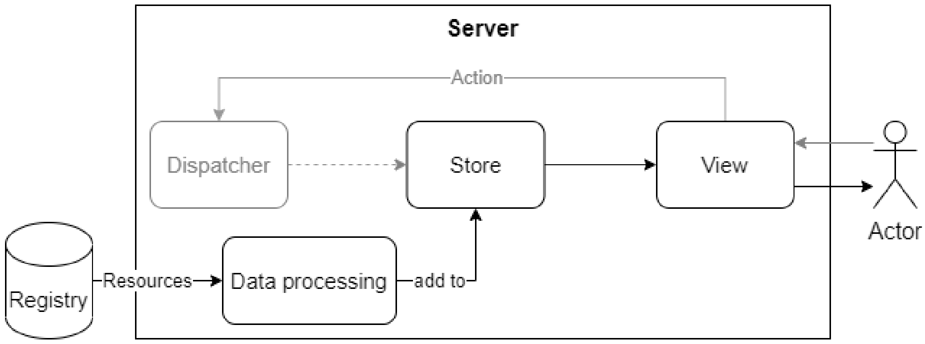


Figure 3.1: Black arrows show the data flow between the high-level components in the system. Grey arrows represent other parts of the system.

controller, receives actions. These events cause certain parts of the application state to recompute, again causing the views to update. Data are not re-fetched from the dental implant registry in these cases.

The data retrieved from the registry and processed and stored on the server are an intricate part of the dashboard, determining much of the dashboard's behaviour and charts. In addition, the data are highly complex, with around fifty variables. Consequently, we need adequate test data to involve users through demonstrations and evaluations when developing such a data-intensive system. The goal was to fulfil this need by creating the dashboard and a synthetic data set in parallel.

3.2 Development Process

Previously, we have completed two of the activities defined by Peffers et al. (2007), which involved identifying problems and solutions (Steps 1 & 2). Next, we will describe the development process consisting of the design science activities *design and development*, *demonstration*, and *evaluation* (Peffers et al., 2007). These are the main activities we used to achieve our goals. However, we have divided them into additional steps to create a more transparent development process.

Figure 3.2 gives an overview of the process in the context of design science. We start each iteration by defining the tasks and evaluation criteria that lay the foundation for developing the dashboard and synthetic data. Subsequently, we create the artefacts in parallel. During the development phase, it was natural for us to investigate the knowledge base, researching aspects such as design, usability, and variable ranges. This knowledge guided the implementation of both artefacts. When we finished the development phase, we conducted an internal evaluation. In Figure 3.2, the lines representing each artefact merge, indicating that we conducted the internal review by combining the two artefacts. This review decided whether we would iterate and add new tasks and evaluation criteria or introduce the artefacts to the environment. If the artefacts upheld a certain quality, we would introduce them to our focus group for demonstration and evaluation. The feedback would then determine the next iteration's tasks

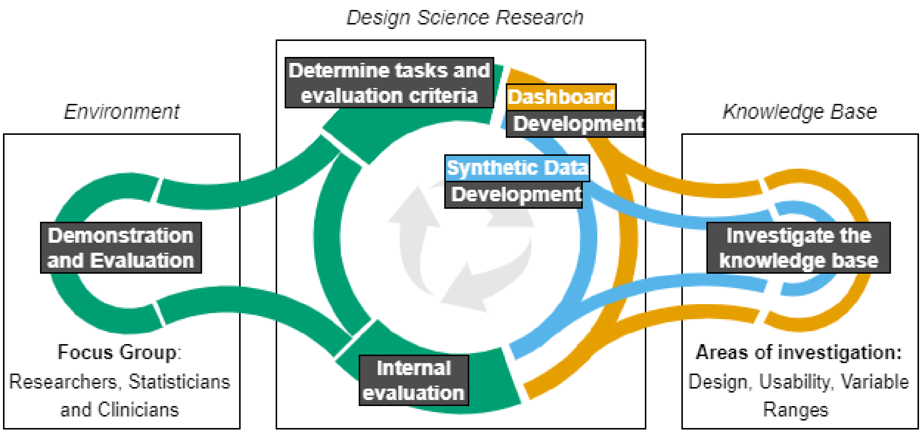


Figure 3.2: An overview of the activities we used in the development process

and evaluation criteria.

In essence, this process allows us to build a high-fidelity prototype populated with a synthetic data set. Consequently, the artefact we produce is not a production-ready application since we will omit several vital aspects, such as security. However, one of our goals is that the prototype may act as a starting point for a proper application. This approach aligns with the activity of evolutionary prototyping. The iterative aspects will also allow us to investigate how the synthetic data and their realism influence the development process. In addition, we may learn of other benefits of producing the artefacts in parallel. The following sections contain a more detailed description of the different steps used in the development process.

Task and evaluation criteria

At the beginning of an iteration, we created tasks to guide the development process. We chose assignments based on the feedback collected from stakeholders through the *demonstration* and *evaluation* activities, which is a part of the *relevance cycle*. However, we did not formulate tasks solely based on user feedback because stakeholders may not know exactly what they want (Parnas and Clements, 1986). A quote from Henry Ford captures this sentiment: "If I had asked people what they wanted, they would have said faster horses". Therefore, we formulated the tasks to explore additional possibilities alongside user feedback. In general, the tasks associated with the synthetic data were closely related to advances in the dashboard, facilitating parallel development.

To understand how the dashboard and the synthetic data set evolved each iteration and potentially affected each other, we summarised the information in figures. Figure 3.3 shows the template we used, populated with a few examples. Here, we display the two artefacts' progress, where blue indicates what functionality we have implemented, and yellow shows the planned improvement for the current iteration.

For the synthetic data in Figure 3.3, a position further to the right on the axis

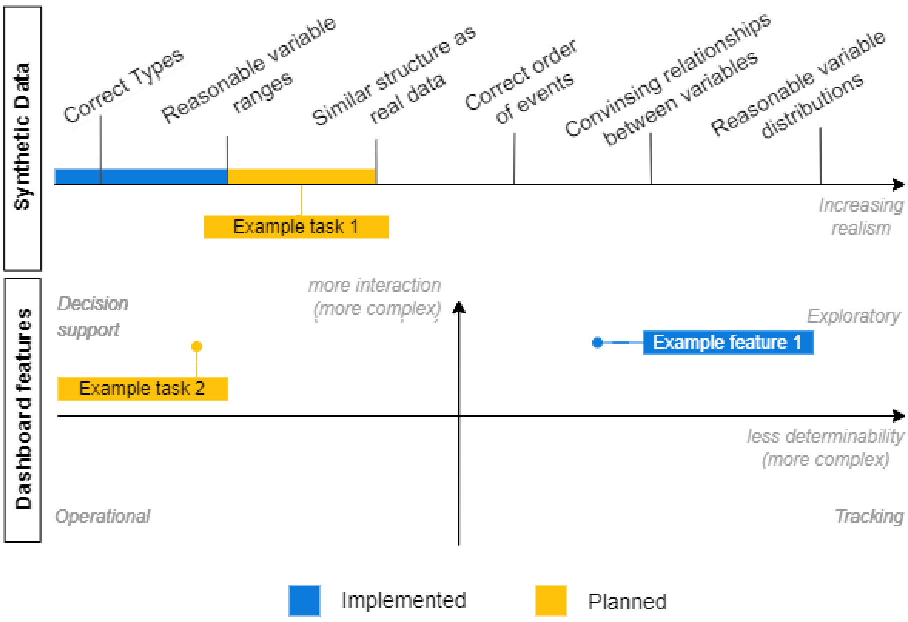


Figure 3.3: We used this template to overview completed and new tasks each iteration. We would place tasks on the axis for the synthetic data set according to how they increased the data's degree of realism. For the dashboard features, we would place features in the coordinate system according to the outcomes determinability and required user interaction. This placement was approximate and indicated the complexity of a feature.

suggests a higher degree of realism. The axis ticks highlight possible actions to improve this attribute. We considered the elements on the left to be easier to implement, but other researchers might order them differently. In addition, the order of the elements may be influenced by the domain, the expertise of the domain experts involved, and the amount of related literature. The equivalent of successfully implementing all steps is creating data that could substitute for real data. Hence, it was impossible to achieve this level of realism using a manual approach. However, the scale served as an appropriate indicator to highlight attributes present in realistic data and to measure our progress. The steps we used were as follows:

- **Correct types -** The synthetic data have the same types as the model it mimics. Hence, developers may insert it into a database created with the data model.
- **Reasonable variable ranges -** The range of variables, such as factors and numeric variables, is within reasonable limits. Consequently, it is challenging to separate single synthetic variables from the actual data.
- **Similar structure as real data -** The conditions influencing whether nullable variables are null or have a value are known and are accounted for in the synthetic data.
- **Correct order of events -** Events should occur chronologically, and the time between them should be sensible.
- **Convincing relationship between variables -** All relationships between variables and their influence on other variables are implemented.
- **Reasonable variable distribution -** The distribution of individual variables should be accurate.

For the dashboard features in Figure 3.3, we used the categorisation suggested by Zhuang et al. (2020), placing them in a coordinate system. The x-axis represents the determinability of the results and the y-axis represents the user interaction. Elements further to the right have lower determinability, and higher-up elements may require more user interaction. For example, the exploratory features were set in the upper right corner of the coordinate system, suggesting high complexity due to low outcome determinability and high user interaction. Our placement of the features in the coordinate system is approximate and is only used to indicate their complexity. In addition, a feature could contain a combination of functionality categories. However, we will only specify one category that best describes the overall functionality.

After establishing the iteration's tasks, we considered which evaluation methods to employ. These methods would assess whether the artefact satisfied the iteration's goals and indicate future modifications as we progressed. Specifying the start and end of an iteration ensured that both artefacts advanced evenly, potentially benefiting from each other.

Development process and the knowledge base

After determining the tasks and evaluation criteria, we started two development cycles, one for the synthetic data set and the dashboard. These cycles were part of the internal *design cycle*, which involved moving from objectives to a design (Peffers et al., 2007). Furthermore, we used the *rigour cycle* in this process, drawing on previous knowledge, theories, and methods grounded in the knowledge base. This information supported the design of our artefacts. For example, we could gain insight into the realistic ranges of the variables or identify solutions to technical aspects of the dashboard.

In the development process, we generated a synthetic data set to fit the data model created by the previous students, defining the distributions and their relationships manually. Most of the iterations contributed to increasing the realism of the synthetic data set, which we highlight in Figure 3.3. We documented the modifications and methods we used to make the changes. Also, we reported any findings or experiences we had when applying changes to the synthetic data. Parallelly, we developed the dashboard according to the identified tasks. Some tasks were precise, while others were vaguer. The less clear goals prompted exploration to try and identify potential solutions, for example, through the *rigour cycle*.

Internal evaluation

We would regularly evaluate the synthetic data set and the dashboard prior to any external demonstration or evaluation. This review ensured that the system maintained a certain level of quality before being introduced to the environment. We considered different attributes depending on the artefact:

- **Synthetic data -** In general, we used automated scripts to assess the tasks associated with the left side of the axis in Figure 3.3. Domain experts will likely be required to assess the right side.
- **The dashboard -** To evaluate the dashboard, we demonstrated and discussed the features in our working group. Our discussions revolved around existing designs and functionality, but also possible additions.

Demonstration and Evaluation

After passing the internal evaluation, we introduced the artefacts for our focus group, which consisted of the following:

- Two statisticians. One is working with another quality register. The other works with a range of Norwegian registries, where he performs tasks such as conducting analyses and generating figures in *R*.
- Two researchers whose research areas are dental implant registries.
- A clinician who is an oral surgeon with excellent domain knowledge related to dental implants.

Not all members were present at each meeting. We demonstrated the dashboard using synthetic data to the focus group, which evaluated its features. This process was part of the *relevance cycle*, where artefacts are introduced into the environment and field-tested.

We used the system implementation scenario described by Zhuang et al. (2020) when evaluating the dashboard. This scenario focuses on stakeholders' perspectives on the functionality offered by the dashboard. Due to the time constraints of the meetings, we used an unstructured format. According to Seaman (1999), this interview format can help extract as much information as possible in a short time. Furthermore, Morgan (1997) found that less structured groups are helpful for exploratory research when you want to learn something new from the participants. Therefore, such an approach was suitable for our research objectives.

We started the focus group session with a demonstration of the dashboard's functionality. For example, we showed participants specific features or use-cases. During the presentation and afterwards, participants voiced their opinions, which we recorded and summarised. It is important to note that when participants reviewed the dashboard, they indirectly reviewed the synthetic data set in the process. This indirect assessment could highlight the need for changes to the synthetic data and potentially uncover errors or misconceptions in the data or the underlying data model.

The information gathered during this activity influenced whether to iterate back or continue with a final assessment. When we completed the project, we performed a more comprehensive evaluation. See Chapter 6 for the final review.

Chapter 4

Design

We present the artefact and its design and development in this chapter, detailing the evolution of the dashboard and the synthetic data set through each iteration. We describe the development process in more detail in Section 1.3.2.

4.1 Iteration 1

In the first iteration, we had to resolve several technical aspects and implement the basic functionality to prepare the artefacts for demonstration and further development. We used a pilot quality registry developed by Vågenes (2022) to establish our foundation. Initially, the database contained small amounts of data, such as standard procedures and materials. However, the registry's data model was the most valuable component, as it was a prerequisite for building our two artefacts. Using this data model as a foundation, we decided on three tasks for this iteration shown in table 4.1.

As a side note, the two evaluation methods shown in Table 4.1 are part of different research cycles. First, the evaluation of the synthetic data is part of the internal *design cycle*, controlling whether the synthetic data maintained a certain quality. Second, the dashboard assessment is part of the *relevance cycle*, introducing the artefact into the environment. We indirectly evaluated

Table 4.1: The tasks and evaluation criteria for iteration 1

	Artefact	**Description**	**Evaluation**
1	Synthetic data	Populate the existing database with synthetic data	Valid SQL-query
2	Synthetic data	Establish some reasonable variable ranges	Focus group
3	Dashboard	Create a dashboard to display part of the data stored in the database	Focus group

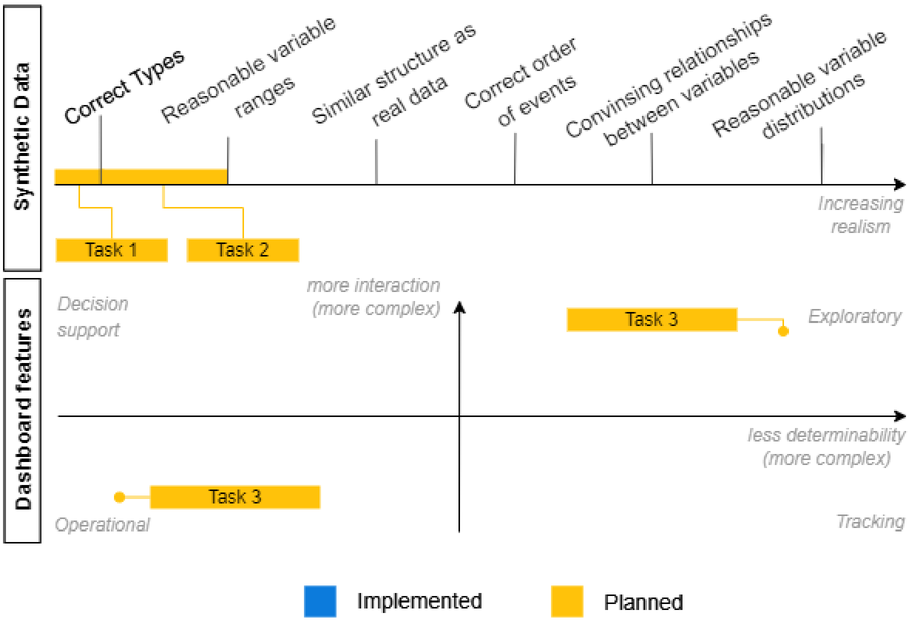


Figure 4.1: Summary of the planned tasks in iteration one and how they affected each artefact.

the synthetic data set in the dashboard assessment.

Task 1-2: Generating data

The method used to generate data in this iteration and solve task 1 was simplistic. We filled each table with random values that fit the data type. Many of the tables without foreign keys already contained data such as standard procedures and standard materials, while many of the tables with foreign keys were empty. Thus, we randomly sampled the primary keys on the owning side for tables that needed foreign keys. Consequently, the generated data were valid from the perspective of the database but otherwise likely neither realistic nor logically valid.

Before inserting the data, we ensured that the data were within reasonable limits. Because the database already contained some standard definitions, the variables that depended on these were guaranteed valid ranges. Other values, such as dates and Booleans, were also simple to create. However, it was more challenging to determine reasonable values of the numeric variables. Therefore, we used the data reported in Raikar et al. (2017), Ahmad and Saad (2012) and Chuang et al. (2002) to gain insights into possible values.

When generating data and inserting them into the database, we detected some poor design choices in the data model. Because a specific implant has many different versions associated with it (denoted "Lot Nr"), it is necessary to use a join table between the implant and the insertion table. This join table stores the "Lot Nr". However, the data model developed by previous students also stored implant properties, such as length and diameter, in this join table. These

Table 4.2: The summarised feedback/results from iteration 1

Artefact	Format	Feedback/Result
Dashboard	Focus group	The current functionality could benefit researchers and registry owners.
Dashboard	Focus group	The dashboard need to distinguish between insertions and removals.
Synthetic data	SQL-query	The synthetic data was successfully inserted into the database
Synthetic data	Focus group	There were some inaccuracies in the synthetic data, and the group had to clarify aspects of the data model

properties should be stored in a separate table to fulfil the second normal form (2NF). We detected these flaws because they made generating synthetic data more complicated than expected. Of course, a detailed review of the database model might have also discovered them. Still, the previous developers did not find these weaknesses, indicating that developing synthetic data might help detect problems.

Task 3: Dashboard development

In addition to populating the database, we created an artefact to communicate and display the data from the database. Through our investigation of the knowledge base, we identified some key elements to guide the development. The dashboard should provide fast and reliable information to stakeholders and should be highly dynamic (e.g., Santos and Eriksson (2014) and Ghazisaeidi et al. (2015)). We used *R* and the *Shiny* framework to implement these fundamental aspects, creating an interactive and flexible application.

The application was a single-page dashboard with a chart and two infoboxes. The chart has a relatively high complexity, allowing users to select the y-axis, the x-axis, and the faceting row (Figure 4.2). This functionality was motivated by studies such as Dagliati et al. (2018), where users could also investigate different aspects of the data. The two infoboxes summarised the data fetched from our database, reporting complication percentage and number of insertions. Studies such as Stattin et al. (2016), where reporting quality indicators positively affected care, inspired these info-boxes. Although we could not define the information displayed in the infoboxes as quality indicators, it called attention to this way of communicating summarised information.

Demonstration and evaluation

We performed the initial demonstration of the dashboard with the group consisting of two scientists and the statistician working with another quality registry. We presented the exploratory features of the current dashboard. We summarised the evaluation in table 4.2

The group agreed that stakeholders such as researchers or registry administra-

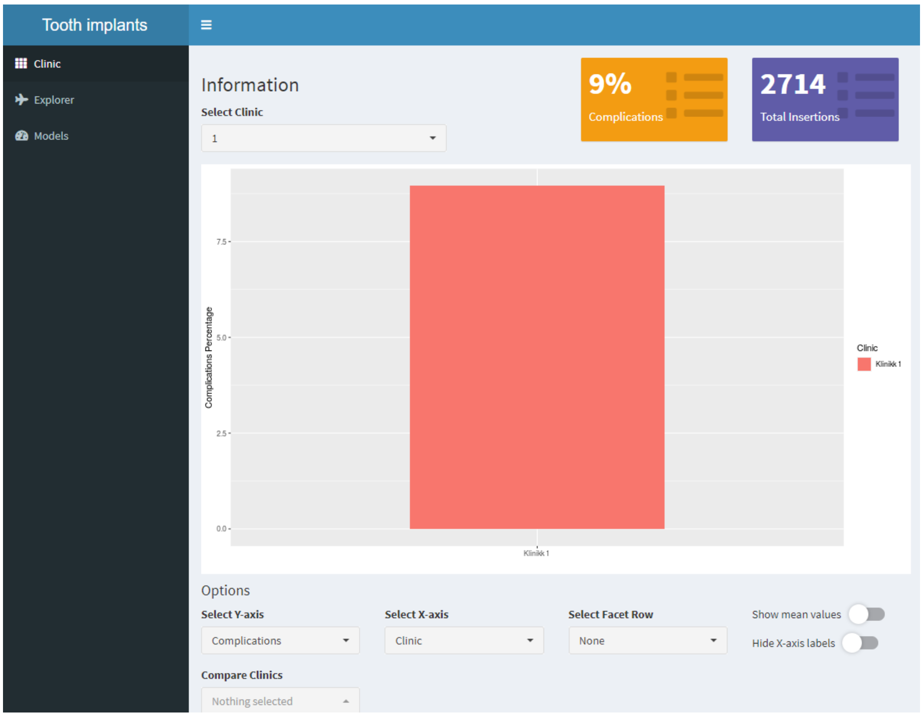


Figure 4.2: Here are the application's features and layout at the end of iteration 1. Already the application has some advanced features, where the user can select the x-axis, y-axis and facet rows and see the result displayed as a graph.

Table 4.3: The tasks and evaluation criteria for iteration 2

	Artefact	Description	Evaluation
1	Synthetic data	Improve validity of synthetic data	Compare against the structure of data submitted to the registry
2	Dashboard	Add a graph which communicates which implants are under-performing	Focus group
3	Dashboard	Add additional features that may be useful for the identified stakeholders	Focus group
4	Dashboard	Improve the layout	Focus group

tors could use the dashboard to explore the data from a dental implant registry. However, it was not particularly informative in its current state. The group suggested functionality to view insertions and removals separately.

Although the SQL query succeeded in populating the database with synthetic data, adhering to the database's constraints, the group recognised some errors during the demonstration. To inform whether an implant had failed, we used the complication percentage. However, this relationship was not logical because complications are only related to insertions, not implying that an implant failed later. Hence, we had to separate the data of the insertions and removals. Furthermore, the group mentioned that the implant's removal reason was better suited to identify underperforming implants.

Due to these misconceptions, the discussion shifted to the attributes of the synthetic data. The group used much of the remaining interview time to clarify other aspects of the data model. This discussion left no time to examine other potential features and new dashboard functionality. However, the group noted that it was an exciting and potentially useful way to summarise data from a registry.

In conclusion, the potential usefulness of our dashboard was already perceptible, but it required more adjustments to create an adequate implementation. This work involved improving the synthetic data set for improved feedback and correcting errors in the application.

4.2 Iteration 2

Based on the feedback from the previous iteration, we established some goals and tasks for iteration 2. We have listed these in Table 4.3.

Although there is only one task related to the synthetic data, its improvement was our main concern as it possibly could contribute to more helpful dashboard feedback. Therefore, instead of generating data based on the database model, it was preferable to change data generation to create data similar to the data submitted by dentists when recording patient information. Therefore, we had to consider the dependencies between variables, identifying when to insert null

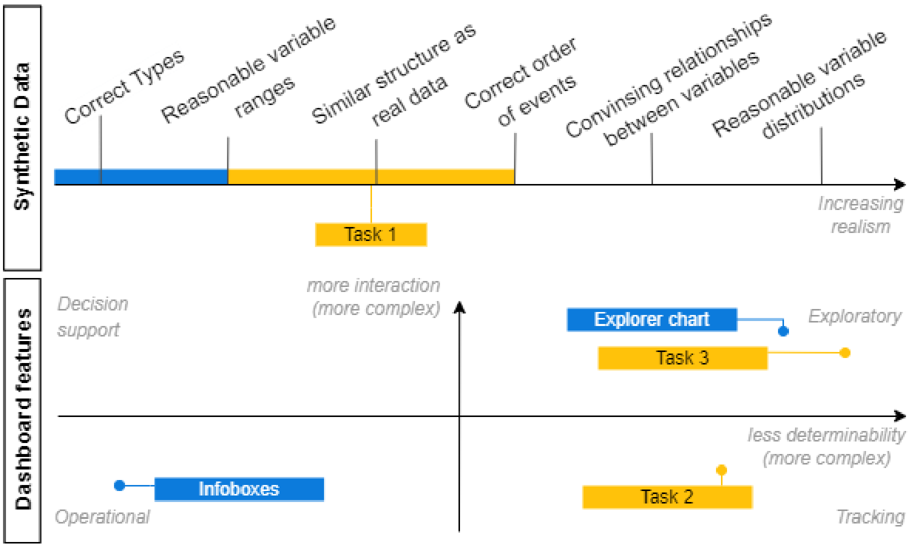


Figure 4.3: Summary of the planned tasks in iteration two and how they affected each artefact.

or a value for nullable variables. Like the previous iteration, we evaluated this goal as a part of the internal *design cycle*. This time a function controlled if the synthetic data conformed to the form submitted by dentists.

The three other objectives were related to the dashboard. They involved several additions and improvements to the artefact. To obtain another perspective, we performed the evaluation with a statistician working with various Norwegian registries.

Task 1: Synthetic data improvements

When generating the synthetic data set in iteration 1, we used a straightforward approach that populated the existing tables with random values. However, we had to consider the dependencies between variables to create a realistic data structure and ensure that the events occurred chronologically. To reduce the complexity of managing this, we had to change strategy. Instead of generating data for each table, we had to create a single table containing all the data before separating it and adding it to the corresponding database tables.

To generate a single table with synthetic data and account for the dependencies between variables, we examined the dental implant registry developed by Fiskeseth (2022) and Vågenes (2022). In particular, we investigated the data submitted to the registry and the grouping of variables in the database model. This investigation allowed us to discern which variables depended on each other, and we structured the data accordingly in a single table. In addition, we ensured that events occurred chronologically, such as the insertion and removal of an implant.

Using this approach, we found that the database design did not reflect the relationship between certain variables in some cases. For example, the database model had associated specific attributes that were only relevant for implant insertions with implant removals. These flaws increased the complexity of generating synthetic data because we had to unravel inconsistencies when comparing the registration application and the database model.

Task 2-4: Dashboard development

Task 2: Add a graph that communicates which implants are under-performing. To accommodate the feedback received on the previous phase's application build, we first added a page displaying the dental implants' removal reasons (Figure 4.4). This feature allowed scientists or dentists to view the implants and why patients had them removed, potentially revealing implants more prone to particular failures. Using the graph's options, the dashboard's users can drill down on removal reasons or implants and compare the performance of different serial numbers of a specific implant. In addition, we added more functionality to the exploratory graph tool (Figure 4.5). We gave the user the option to view data related to insertions or removals, specific clinics, or combine all data.

Task 3: Add additional features that may be useful for the identified stakeholders. Furthermore, we expanded the prototype with new features that the focus group did not request, but which we considered potentially valuable. During the previous iteration's evaluation, part of the discussion was focused on the relevant users. The group's opinion was that researchers and registry administrators suited this type of dashboard. Therefore, we chose to adapt it further along the scientific axis, taking this into account. We added analysis capabilities to enable users to create models and test for statistical significance, with support for linear models and binomial logistic regression (Figure 4.6). Compared to using plain *R*, our approach made creating a model more accessible because it wrapped the functionality in a convenient UI. Similar observations were made by Zhou et al. (2020) when developing a comparable application in health care.

Task 4: Improve the layout. The charts created by the application may include many factors and facet rows, requiring a lot of space to be readable. Thus, the current options placement was not ideal, taking up much real estate at the bottom of the screen. We identified comparable studies, such as Marini and Binder (2017) and Soneson et al. (2020), solving this differently by positioning all user inputs in a sidebar. We adopted this design choice, giving more space for larger graphs.

In addition to new features and functionality, we reduced the technical debt of our codebase. There had been an accumulation of technical debt in iteration one because we were unfamiliar with the *shiny* framework. At one point, we had written much of the application in one large file, and it was difficult to make changes. We cancelled this debt by separating each feature into different components. Consequently, expanding the functionality and adding new features became easier.

Demonstration and evaluation

Our primary goal in this iteration was to improve the validity of the synthetic data set. We engaged the statistician working with various Norwegian registries.

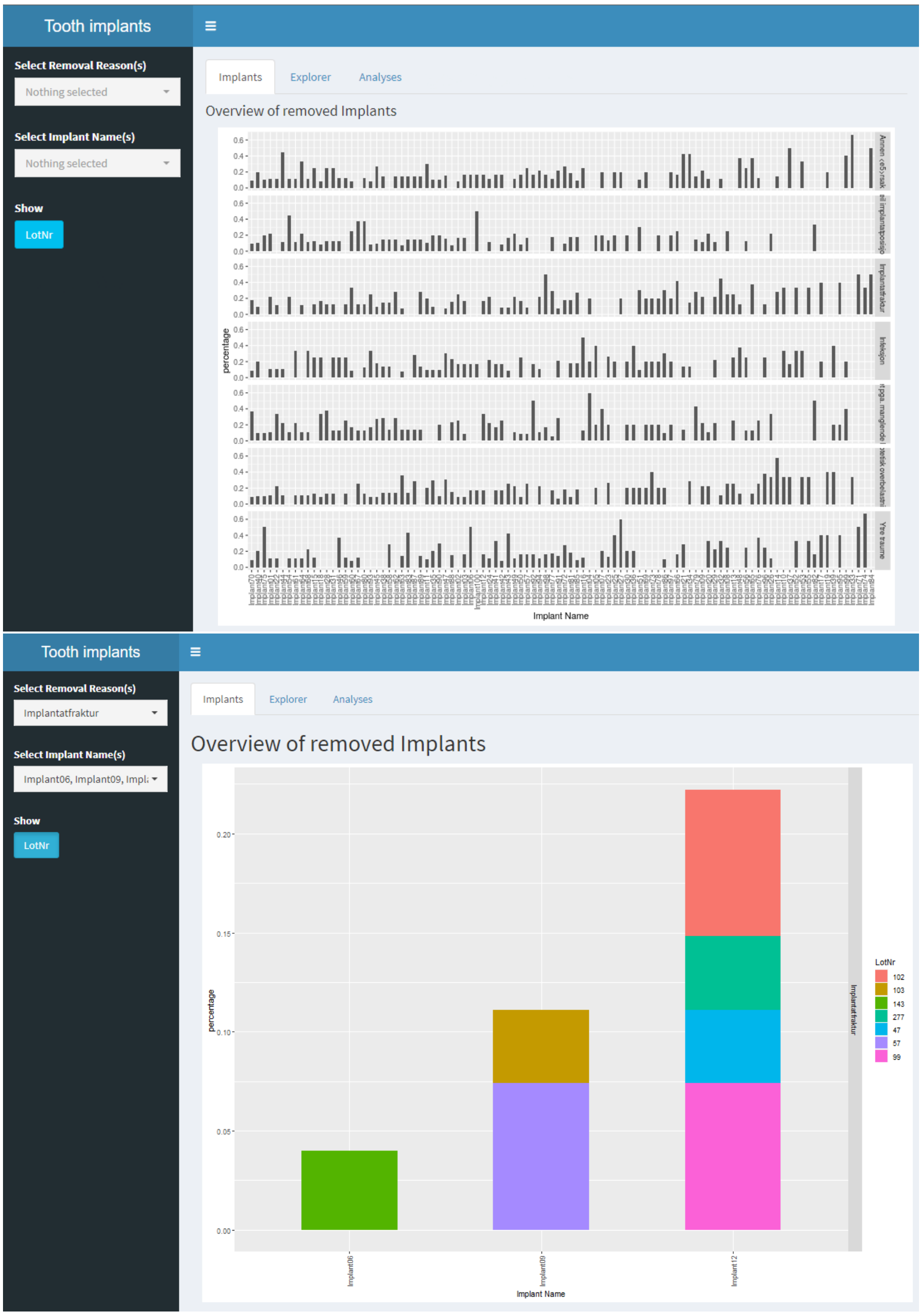


Figure 4.4: The top image shows the landing page of the dashboard in iteration 2. Both graphs display the removal reason percentage of an implant on the y-axis, implant names along the x-axis and the removal reasons in the facet rows. In the sidebar, the user can drill down on particular removal reasons and implants and even show the serial numbers of specific implants. The second image shows an example of this selection.

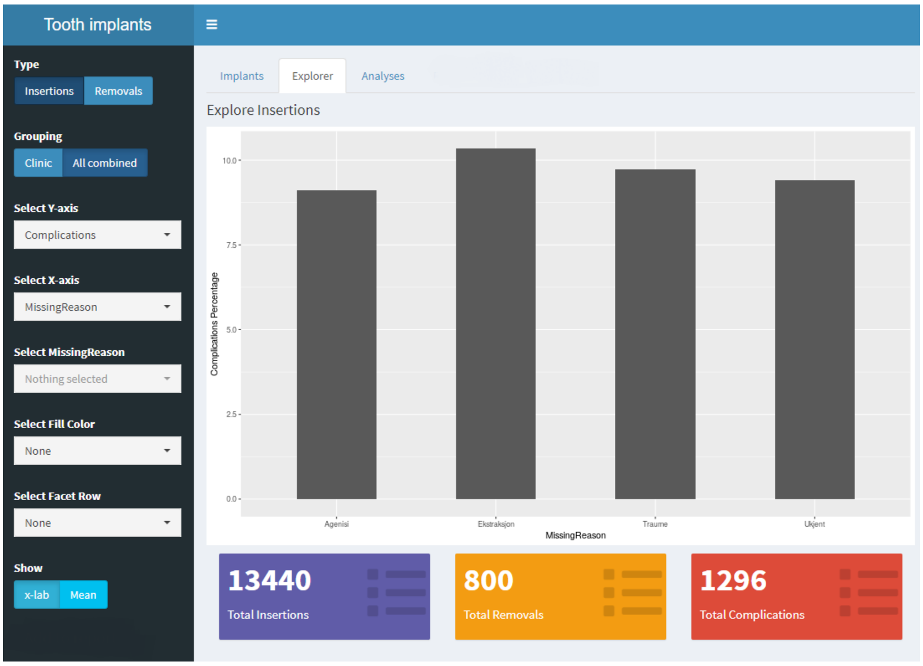


Figure 4.5: Here is the exploratory part of the dashboard from iteration 2. Since iteration 1, we have given the tool increased functionality and features. Now users can select removals or insertions, focus on particular clinics, or combine all the data.

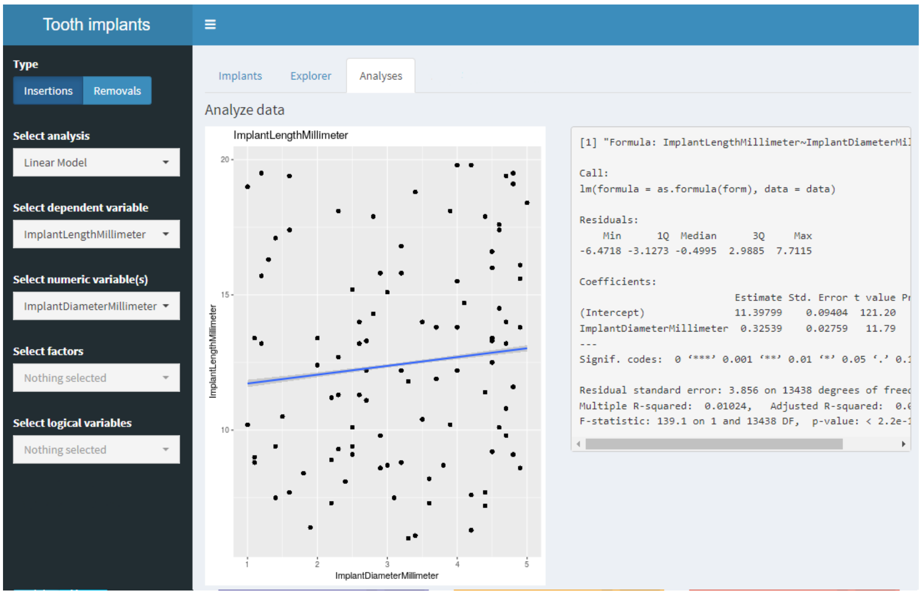


Figure 4.6: Here is the analysis part of the dashboard. The user can select the model type and dependent and independent variables. If possible, the dashboard will display a plot alongside the summary statistics of the model.

Table 4.4: The summarised feedback/results from iteration 2

Artefact	Format	Feedback/Result
Synthetic data	R-script	The synthetic data have a similar structure as the data submitted by dentists.
Dashboard	Focus group	The exploratory features of the dashboard could benefit registry administrators and clinics
Dashboard	Focus group	The graphs are cluttered and not optimised for readability
Dashboard	Focus group	The dashboard could benefit from dynamic reports.

He was familiar with the technical aspects of our application but not the specific domain. We summarise the feedback in Table 4.4.

To test the synthetic data set, we employed an R-script to confirm that its structure was similar to the information registered by dentists. The script verified that the variables were only present if their dependencies were present. Moreover, it confirmed that the dates came naturally and identical implant models had similar attributes.

In response to the short demonstration of the dashboard, the statistician made similar remarks as the scientists and the statistician in the previous iteration, recognising that exploring the data could benefit various stakeholders. Additionally, he mentioned that we had not optimised the dashboard's themes. He identified unnecessary lines and colouring. More importantly, however, he suggested modifying the application to generate dynamic reports. This suggestion was highly relevant for our dashboard, as it could merge the benefits seen in annual reports with the dynamic nature of dashboards.

A dynamic report includes both dynamic and static elements. Often, the static parts are predefined portions of text and the document's structure, while the dynamic components are text, graphs, and tables that depend on data. The amount of fixed and variable content depends on the template. Nevertheless, the changeable content is generated based on the current data in the database. Therefore, stakeholders can quickly generate up to date reports whenever desired and forward them to relevant stakeholders. This feature enables stakeholders to access the registry's data without knowledge of the dashboard's functionality.

In summary, this iteration improved our synthetic data set by structuring it according to data registered by dentists. Furthermore, this iteration indicated that our prototype could benefit from additional features to communicate the information stored in the registry.

4.3 Iteration 3

Since we had improved the realism of the data, we next considered the statistician's suggestion to add reporting functionality. Although the prototype was

Table 4.5: The tasks and evaluation criteria for iteration 3

	Artefact	Description	Evaluation
1	Dashboard	Add report functionality	Focus group
INTERNAL EVALUATION			
2	Dashboard	Reduce the cognitive load of removal reason chart	Focus group
3	Dashboard	Add a graph displaying how a variable changes over time	Focus group
4	Dashboard	Create a uniform style for the different charts	Focus group

already feature-rich, we believed that it was beneficial to expand the functionality further to demonstrate a broader range of its capabilities to stakeholders. We give an overview of the tasks and evaluation methods in Table 4.5.

Implementing the reporting functionality was initially our sole task during this iteration. However, after conducting a comprehensive internal evaluation, we decided to add some jobs. The synthetic data set remained unchanged because it was sufficient for the planned tests.

Task 1: Dashboard development

We began the development by exploring the possibility of generating reports, and we quickly identified an option to use *rmarkdown*. R markdown is a markdown language for creating dynamic documents, combining fixed text and R code (*rmarkdown: Dynamic Documents for R* 2021). We created a report template with static text and included some of our application's functions to achieve similar visualisations. Furthermore, the flexibility of the *R* and *shiny* framework allowed us to integrate the generator with the application, enabling users to download reports. Initially, we supported HTML as the report format. An option we considered, but dismissed as out of scope, was to connect an email service for stakeholders to share the reports. This alternative would ensure a much easier distribution of information. However, we considered the downloadable information sufficient for this project.

We implemented a basic version of a standard and customisable report. To create the customisable document, users could click the "Add to Report" button during their data exploration. This dashboard feature saves the user's current graph, enabling them to combine them in a report. Hence, the document consisted mainly of R-functions and had little fixed text. Alternatively, users could download a predefined report with limited customizability. The similarity between both of these methods was that they used the most up to date data to generate graphs.

Internal evaluation

During the internal evaluation, we inspected and discussed the report generator and the application's design. Overall, we concluded that the artefact's capa-

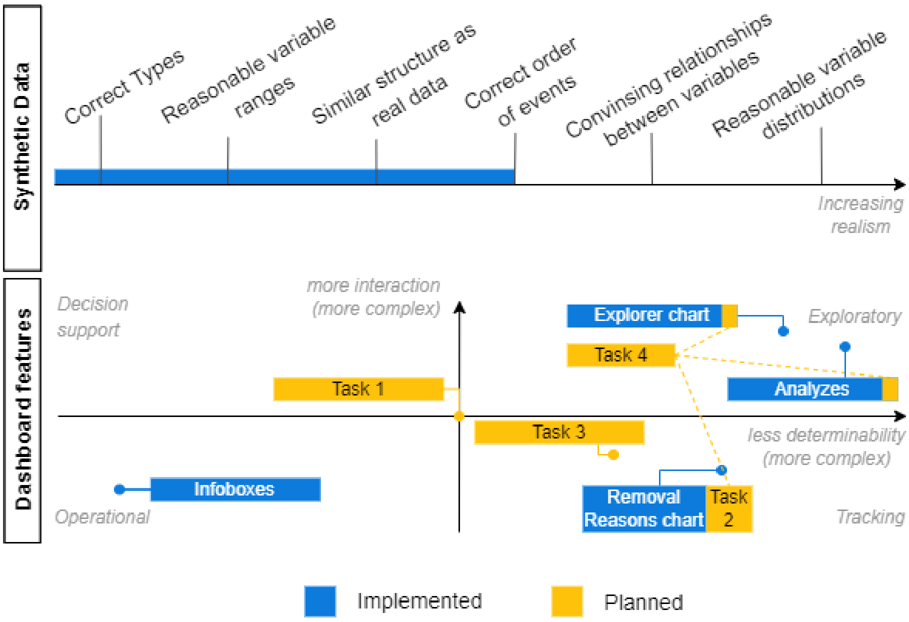


Figure 4.7: Summary of the planned tasks in iteration three and how they affected each artefact.

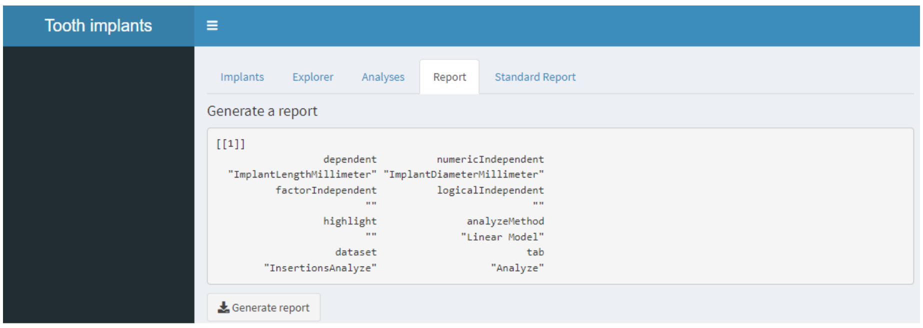


Figure 4.8: Here is the reporting functionality added to the application. We added two new tabs, giving the user options to generate a customisable report and a standard report. In the screenshot, the user can create a custom report and has added one graph. The application displays the information concerning the selected graph as text. The application will use the arguments to create the chart when clicking generate.

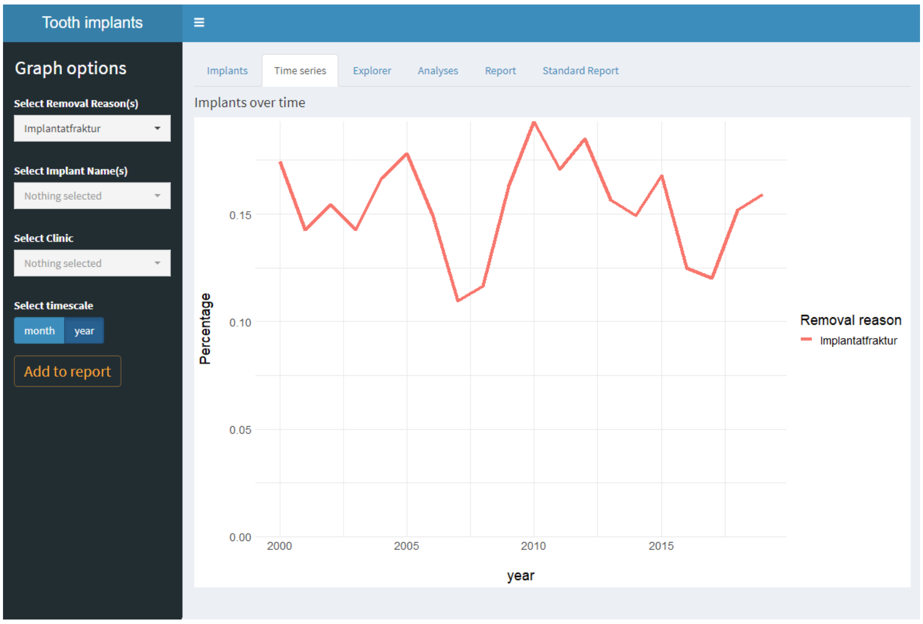


Figure 4.9: Here is the time series chart, giving insights into how implants and clinics change over time. The functionality offered by this chart may support longitudinal studies.

bilities were potentially valuable and engaging. However, we recognised some missing features and challenges with the visualisation. We had to improve these aspects prior to external evaluation. Otherwise, the feedback could be of lower quality.

Task 2. Reduce the cognitive load of the removal reason chart. In figure 4.6, the users are shown an enormous amount of data, with hundreds of bars in a barplot. We assessed that such a display might reduce users' engagement, despite being useful when focusing on particular implants; therefore, we made two changes. First, we simplified the graph to improve users' ability to recall information (Lusk and Kersnick, 1979). This simplification was performed by aggregating the data, focusing on the reasons for the removal of all implants (Figure 4.10). The user can still investigate individual implants, but the initial chart displays a limited amount of data. Second, we flipped the axis because this better suited long labels, as suggested by Wickham and Grolemund (2017) (Ch. 3.9).

Task 3. The prototype had no charts displaying how variables change over time. The missing feature we identified in the internal evaluation was a time series chart. Information changes over time, possibly revealing intriguing trends. Before this iteration, none of our visualisations displayed this information, hiding an aspect of our synthetic data. Therefore, we included this feature to highlight how our dashboard can present changes over time and support longitudinal studies. The graph we produced showed changes in removal reasons over time, and the user could narrow it down to specific causes, clinics, and implants.

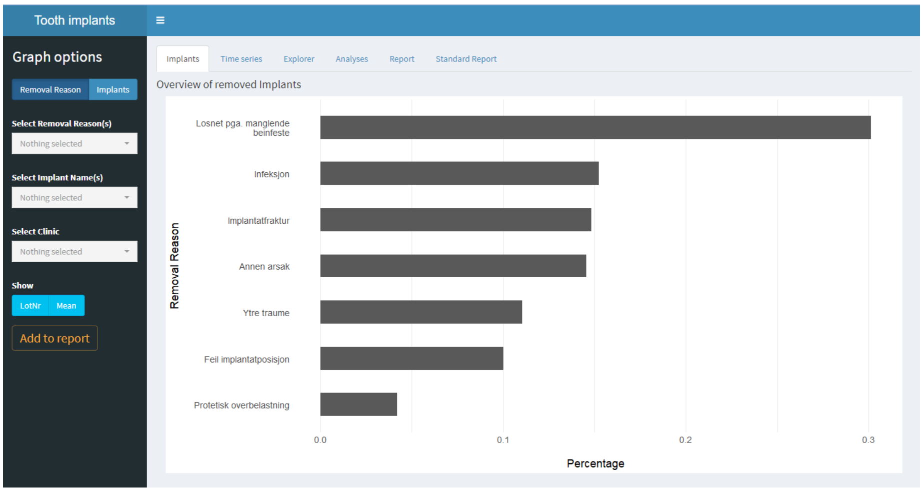


Figure 4.10: Here is the simplified overview of implants. Instead of showing all implants, we aggregated the data. Therefore, the users need to process less information to interpret the chart. In addition, extra clutter is removed, making the presentation more readable.

Task 4. Create a uniform style for the different charts. The statistician first raised this issue regarding the plots' readability in iteration two. To simplify the graphs, we removed anything that did not aid in understanding the data, referred to as "non-data ink" by Tufte (2001). For example, the grey background. In general, we focused on simplification and emphasis when improving visualisations (Evergreen and Metzner, 2013).

Demonstration and evaluation

Because of the improved aspects of the synthetic data set and application, we expected that the feedback in this iteration would be more detailed. It was likely easier for stakeholders to recognise functionality, features, or data that are lacking or need modification. We presented the application to the two scientists and one statistician from iteration 1. As expected, we received more specific feedback during this evaluation. This time, the correctness of the synthetic data received less attention, and the dashboard's features were in focus. We present a summary of the feedback in Figure 4.6.

Previously, we had not made any effort to connect insertions and removals. We generated these events separately and randomly assigned them to a person. However, the evaluation group wanted to explain dental implant removals with factors related to the insertion operation. Therefore, creating a tighter coupling between insertions and removals was necessary for our synthetic data.

Regarding the dashboard, the group appreciated many of the new features. The analysis functionality caused one participant to say to the other participants: "Now, can you also use R". The group agreed that this was an exciting feature that stakeholders could use to find areas worth further exploring. Inspired by this part of the demonstration, they suggested adding a survival analysis feature.

Table 4.6: The summarized feedback/results from iteration 3

Artefact	Format	Feedback/Result
Synthetic data	Focus group	Link insertions and removals by persons.
Dashboard	Focus group	Useful with analysing possibilities to identify trends. Survival analysis would also have been interesting.
Dashboard	Focus group	Report functionality can be useful when wanting to share specific plots. Difficult to use for annual reports because you always want to customize some charts
Dashboard	Focus group	More options to group the data, for example, by how long the implant survived.
Dashboard	Focus group	Use funnel charts instead of bar charts to better judge variance

Furthermore, the group expressed that the time series plot could be helpful, but there was a need to filter the data. They wanted to perform filtration based on survival time. For example, by only reviewing implants removed after the first year, it could be easier to help identify why dentists removed some implants early.

We also discussed the "Explorer" feature. The statistician suggested funnel plots to compare factors with a small number of events with factors with more events. The current bar charts did not communicate this variance. However, adding another visualisation technique to the plot would likely increase the complexity. Instead, a possible solution would be to add more pages handling specific tasks and not have just one visualisation responsible for all functions. The group suggested using annual reports from other quality registers to identify helpful charts.

In summary, the additional features added to the dashboard may be beneficial but still require more tuning to communicate the most relevant information. Interestingly, the current functionality demonstrated with the synthetic data set seems to engage the evaluation group, causing them to suggest new features.

4.4 Iteration 4

The feedback from iteration 3 was substantial and prompted us to make significant changes and add more functionality to the artefacts. We present an overview of the new tasks and evaluation criteria in Figure 4.7.

Enhancing the synthetic data set by linking insertions and removals of a patient was an important goal for this iteration. This change would improve the possibilities for analysing removals in the dashboard due to having access to data from the insertion procedure. Furthermore, we created more specialised pages in the dashboard.

Table 4.7: The tasks and evaluation criteria for iteration 4

	Artefact	Description	Evaluation
1	Synthetic Data	Create patients with related insertion and removals	Control if the processed data in the dashboard has the same relationships as the generated data
2	Synthetic Data	Create implants with lower survival rates to test the survival analysis	Control whether the implants have a lower survival rate using the dashboard's survival analysis
3	Dashboard	Survival analysis	Focus group
4	Dashboard	Filter insertions and removals based on implant survival time	Focus group
5	Dashboard	Funnel plot to show variance	Focus group
6	Dashboard	Move the clinic options from the explorer plot, merging it with time series graph and infoboxes	Focus group

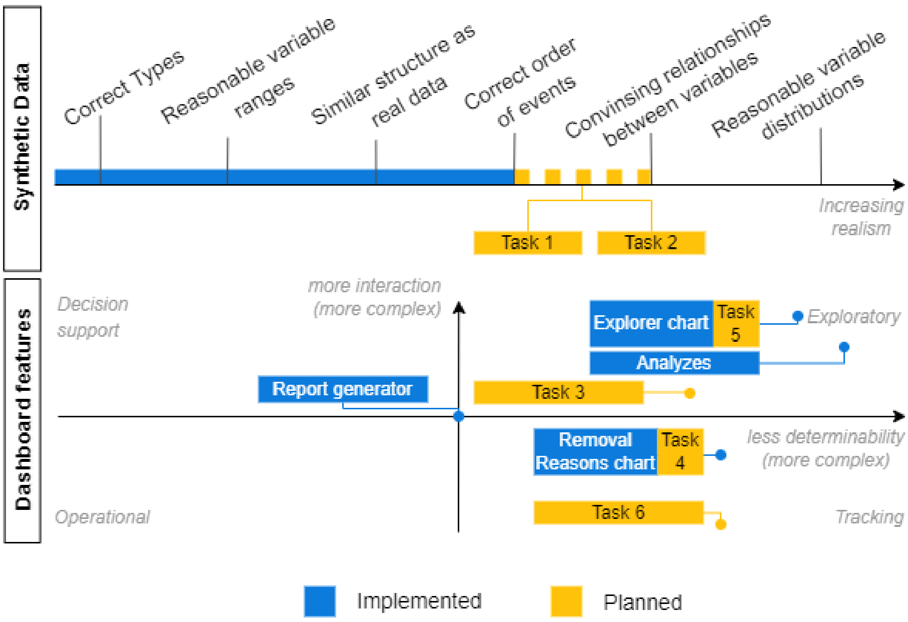


Figure 4.11: Summary of the planned tasks in iteration four and how they affected each artefact.

Task 1-2 Synthetic data improvements

The data generation strategy employed in this iteration was similar to the previous iterations. We produced a large table containing all data of the database model. Previously, we had created insertions and removals separately and randomly assigned them to patient entries in this table, but this approach was no longer viable. We needed to ensure that removals and insertions of patients' dental implants were related. This connection would allow us to join them during the dashboard's data processing.

To ensure such a relationship, we created a table with a given number of patients and iterated through them. We first produced a random number of insertions for each patient and then did the same for the removals. Afterwards, we associated some of the insertions with some of the removals. Therefore, some patients had their implants inserted and removed, or just one of the events. At the same time, we solved task 2, increasing the probability of removals for certain dental implants.

Consequently, we increased the realism of the synthetic data set but only a fraction. We still generated most of the variables in isolation without any interaction with other variables. The task of manually implementing believable relationships between all variables is likely impossible.

Task 3-6 Dashboard development

We first solved *Task 3* using the *survival* package in *R* and displaying the generated model in a graph (Therneau, 2022). Figure 4.12 shows the feature. In addition to this survival analysis, we identified a potentially helpful chart related to implant survival in the annual report of The Norwegian Arthroplasty Register. This chart compared the survival of knee joint replacements after three and ten years (Furnes et al., 2020). Therefore, we developed a similar visualisation for dental implants to supplement the survival analysis (Figure 4.13).

Next, we solved *Task 4* by adding a grouping variable during the server's data processing. This variable categorises implants according to their survival time. Figure 4.14 shows how users could group facets based on this factor.

Implementing a funnel chart in *Task 5* was more challenging to solve due to the already extensive functionality of the "Explorer graph". Consequently, we rewrote and simplified much of the code responsible for this graph's functionality, making it easier to accommodate new additions. Figure 4.15 shows the new chart.

Simultaneously, we moved the clinic options out of the "Explorer graph", solving *task 6*. We show this in Figure 4.16. It is important to note that we have not verified that the information displayed on this page is helpful to clinics. However, this layout and these visualisations indicate a potential method for introducing information to clinics.

Demonstration and evaluation

For this demonstration, the same two scientists and one statistician were present. In addition, a clinician with excellent domain knowledge had also joined. Con-

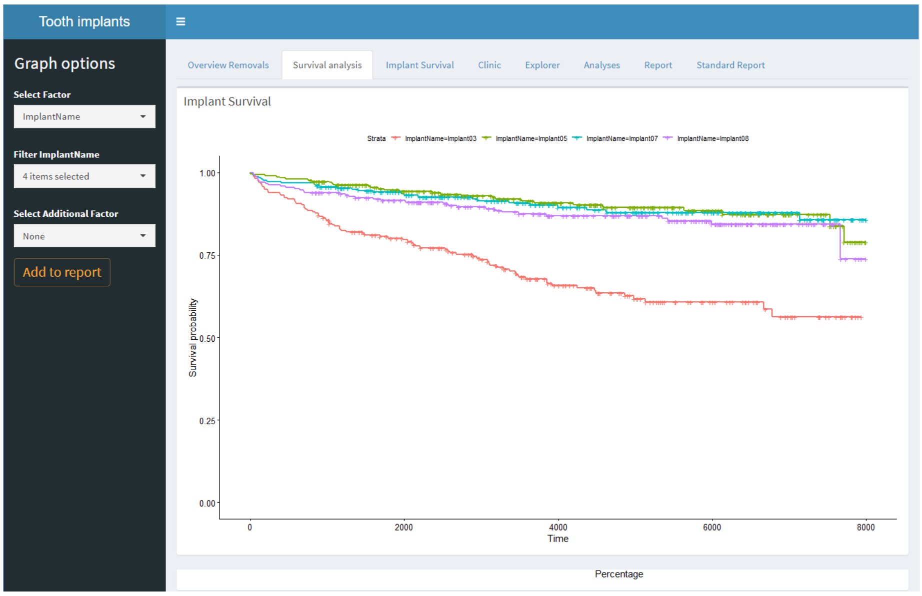


Figure 4.12: Survival analysis of different implants

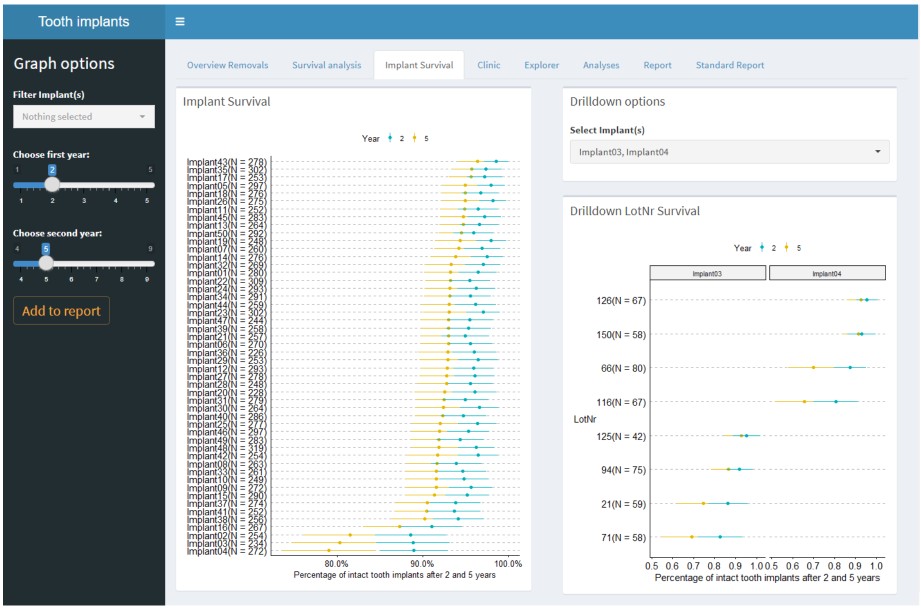


Figure 4.13: The survival rate after two and five years for different implants

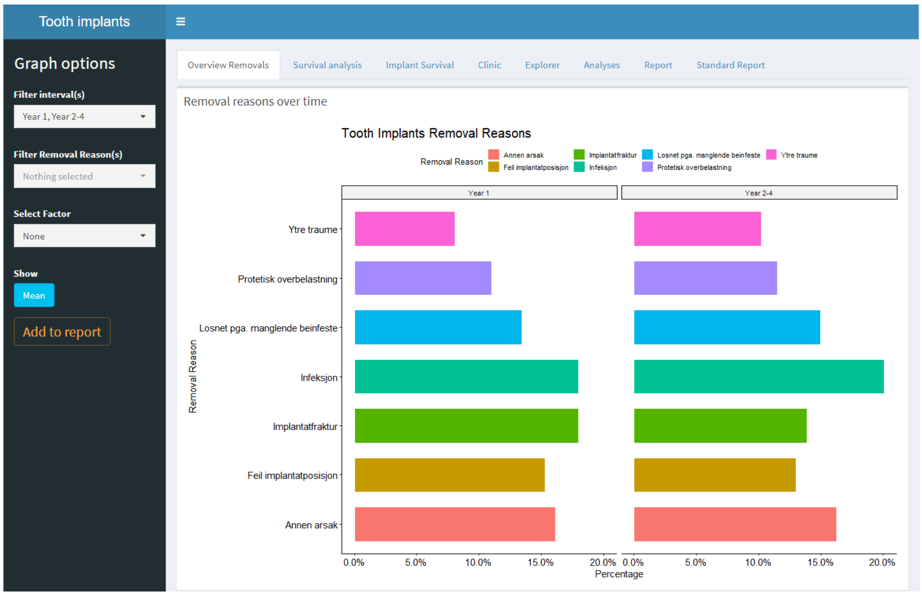


Figure 4.14: Users can now view removal reasons categorised according to survival time in years.

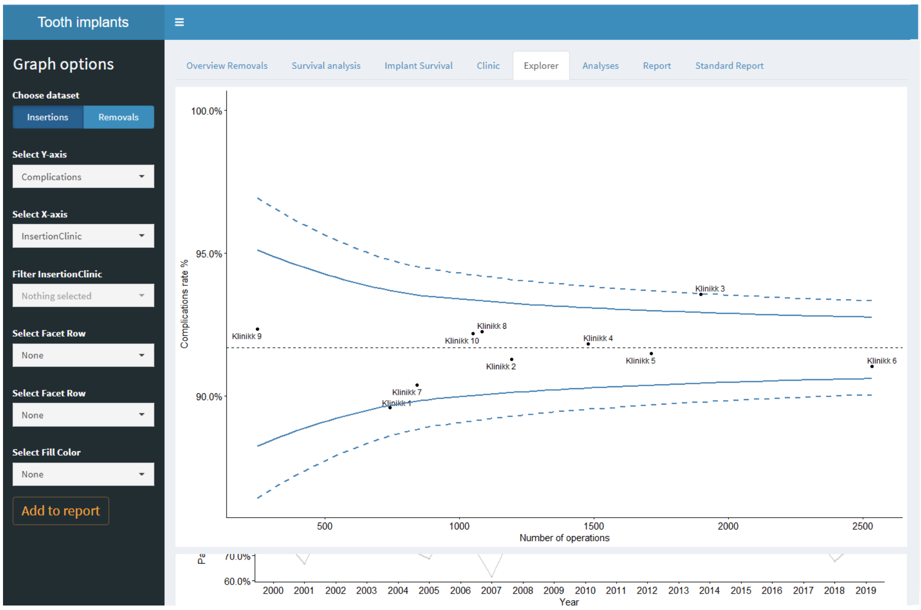


Figure 4.15: The explorer plot supports funnel charts, and we extracted the clinic options to another chart.

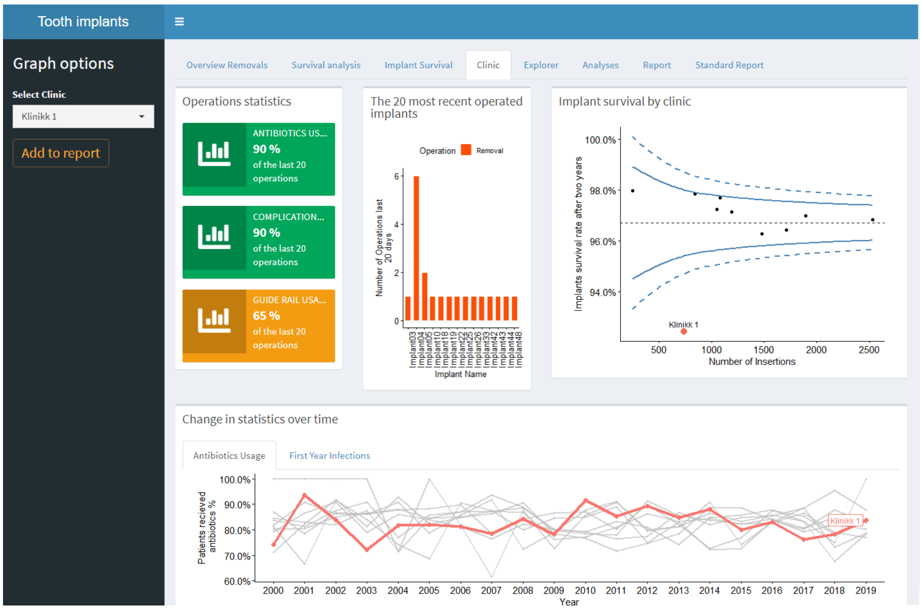


Figure 4.16: Information that potentially is relevant to clinics.

sequently, it was possible to get a perspective on the dashboard's clinical use.

We completed the tasks related to the synthetic data set. First, we did link the patient's insertions and removals in the application. We compared several patients before adding the synthetic data to the database and after retrieving them from the database in the application, and the information was the same. Second, we confirmed that there was a lower survival probability for the implants we had selected to fail more often in our synthetic data generation. These observations indicate that we can create various scenarios using synthetic data and use them in our tests.

The clinician and scientists agreed that the dashboard's charts should focus less on implant names. These names are generally of little interest, except if clinicians should explicitly avoid some implants. Instead, other groupings would be more interesting to explore, for example, implant length.

The clinician also stressed the importance of comparing similar groups of implants, patients, and clinics. Due to the variance within these groups, combining all the data is not informative when exploring the charts. He suggested a type of filter to select patients, implants, or clinics with specific attributes. For example, this would allow users to compare only patients with a particular bone volume.

The focus group also commented on the complexity of the "Explorer" feature. They suggested a more accessible chart to overview the distribution of a variable before freely exploring the data.

To summarise, we demonstrated the potential for introducing more advanced relationships in the synthetic data set and using it to test dashboard features. For the dashboard's charts to be valuable, we must allow users to compare

Table 4.8: The summarised feedback/results from iteration 4

Artefact	Format	Feedback/Result
Synthetic data	Observation	Insertions and removals by persons are identical when comparing the generated data with the data added to the database and processed by the dashboard
Synthetic data	Observation	The implants we created with a higher probability of failure was possible to identify in the survival analysis.
Dashboard	Focus group	All charts should be less focused on implant names; it is often more interesting to group implants by other attributes.
Dashboard	Focus group	Charts should compare patients with similar characteristics to be more informative.
Dashboard	Focus group	The complexity of the explorer graph may be too high. The group suggested separating simpler and more advanced graphs.

Table 4.9: The tasks and evaluation criteria for iteration 5

	Artefact	Description	Evaluation
1	Dashboard	Add a global filter to ensure that similar groups of patients, implants or clinics are presented in the charts.	See final evaluation chapter
2	Dashboard	Reduce the complexity of some of the dashboards options	See final evaluation chapter

similar groups of patients, implants, and clinics.

4.5 Iteration 5

Due to the project's time constraints, this iteration was our last. We wanted to add functionality to improve the charts' usefulness and possibly simplify parts of the dashboard. We decided to merge the evaluation with a final system evaluation, which we present in the next chapter. Table 4.9 shows the iteration's tasks.

Task 1-2: Dashboard development

We added a global filter shared across all dashboard pages. The user can select one factor and choose which levels to include. This filtering allows scientists and clinicians to compare similar data when viewing various charts.

To simplify the layout and usage of the dashboard, we group the options under panels. According to Leung and Cockburn (2020), open panels or collapsable

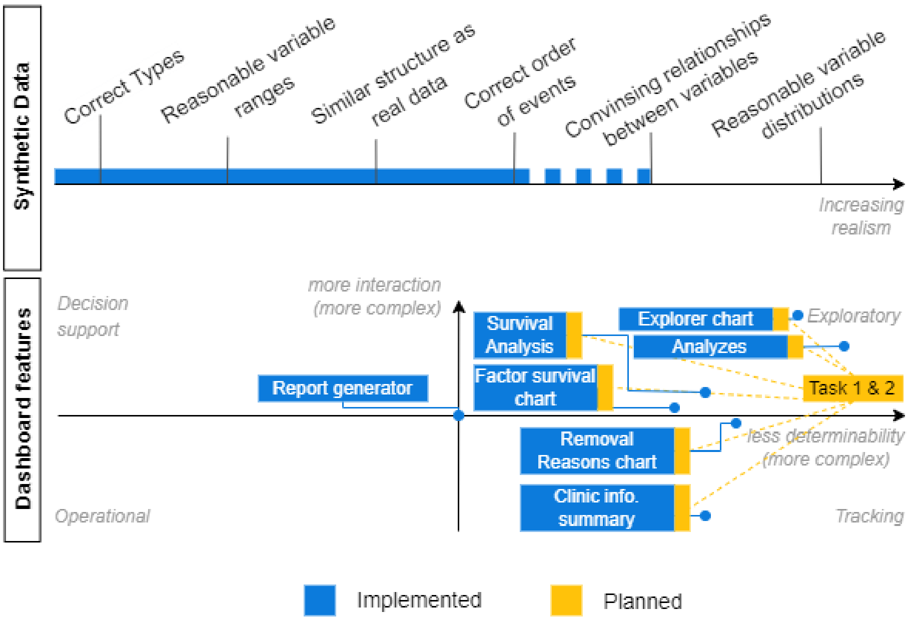


Figure 4.17: Summary of the planned tasks in iteration five and how they affected each artefact.

panels should be preferred by designers if there are few sub-menus. We decided to add three sub-menus: global filter, basic options, and advanced options. Consequently, we chose collapsable panels with basic options starting expanded. This design may direct stakeholders to focus on the most commonly used options and hide options that add complexity. However, the downside of using collapsable panels is that it can lead to longer task times in certain advanced use cases (Leung and Cockburn, 2020). Figure 4.18 shows the new layout when you first land on a page and expand all panels.

We did not choose to separate the graph options under basic and advanced panels for the Explorer chart because of the close relationship between these options. Therefore, we added a switch that activated the advanced options (Figure 4.19). The basic option shows the distributions of various variables, while the advanced option allows users to change the x-axis, y-axis, and grouping.

Lastly, we reduced the focus on implant names in the implant survival chart and gave users control over which factors to explore (see Figure 4.20).

4.6 Section summary

The design phase started with a simple synthetic data set and an application that displays one chart. Throughout the iterations, we modified the synthetic data set, increasing its realism. At the same time, we added and adjusted the dashboard features. A more detailed summary is listed below.

- **Iteration 1** - The synthetic data set fitted the database model and had

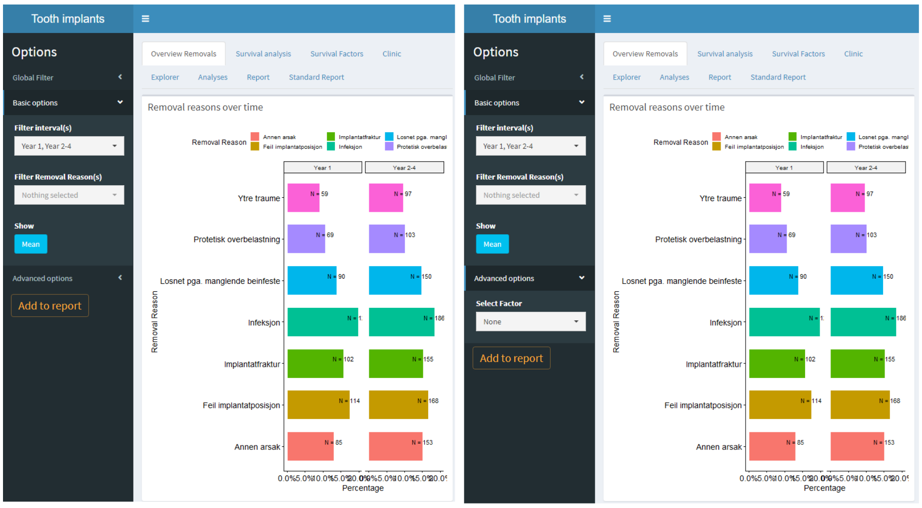


Figure 4.18: Advanced options are hidden by default when redirecting to a new dashboard page.

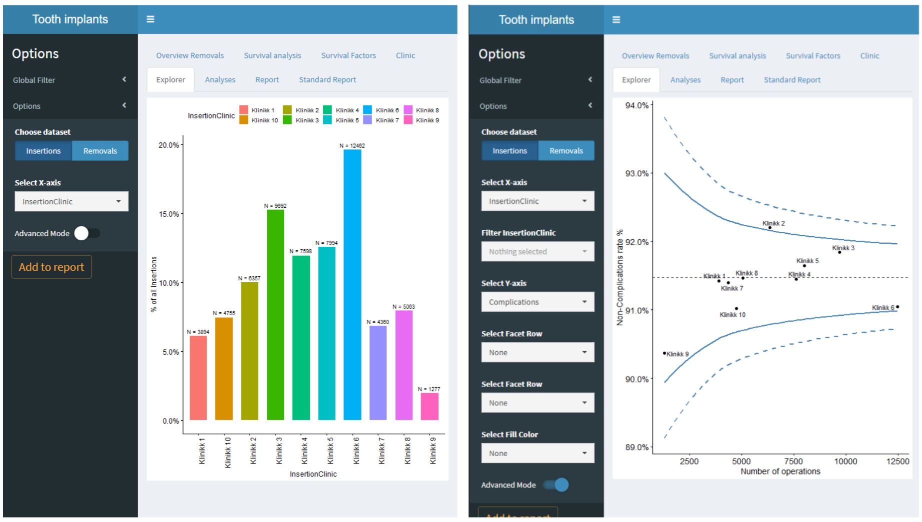


Figure 4.19: The explorer chart has an advanced and basic options switch that changes the chart and the number of available options.

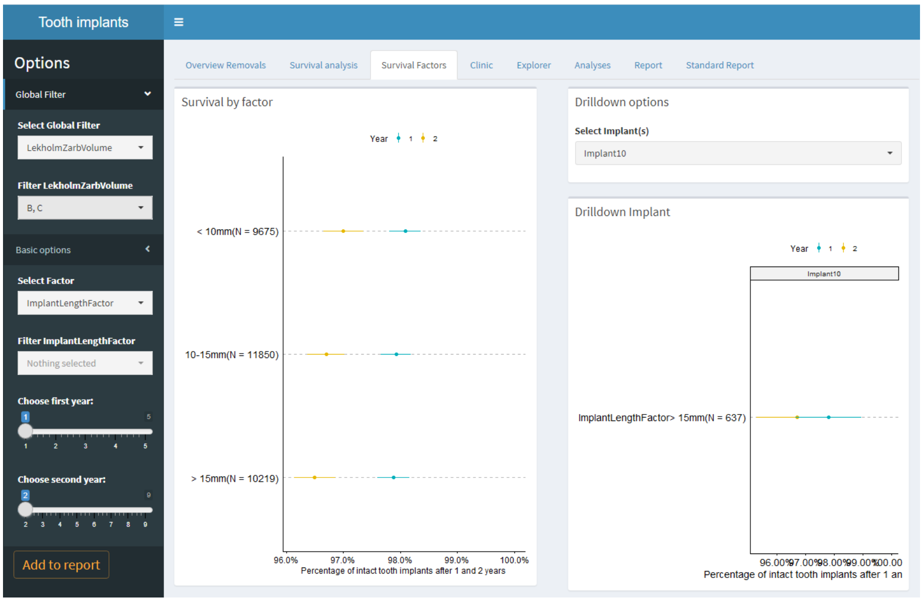


Figure 4.20: Less focus on implant names. This chart allows users to choose a grouping factor to compare survival rates after two time periods.

reasonable variable ranges. The dashboard consisted of an exploratory barchart. We discovered that the data model was not optimal when creating the synthetic data, with some attributes being misplaced. The focus group agreed that the dashboard might be valuable for a visual overview of the data stored in the registry, but it needs to differentiate between insertions and removals. In addition, there were some inaccuracies that we needed to correct in the synthetic data.

- **Iteration 2** - We modified the synthetic data set to have a realistic structure with correct dependencies between variables. Also, we ensured that any events occurred chronologically. We discovered that some variables were not optimally grouped in the database when developing the synthetic data, resulting in many empty fields. The dashboard's exploratory chart was enhanced, and we added analysis capabilities. The focus group assessed that the dashboard could benefit from report functionality and improved chart readability.

- **Iteration 3** - The synthetic data set remained unchanged. The dashboard gained report functionality, a time series chart, and general improvements in readability. The focus group agreed that the analytical capabilities were practical, which led to them suggesting additional features. Additionally, they said that the reporting functionality was helpful for sharing plots but could probably not replace annual reports.

- **Iteration 4** - We improved the synthetic data set by creating patient histories of insertions and removals. We also added a lower survival probability for some implants to test a feature. The dashboard received several changes. We created a survival analysis and a new page for clinics to view

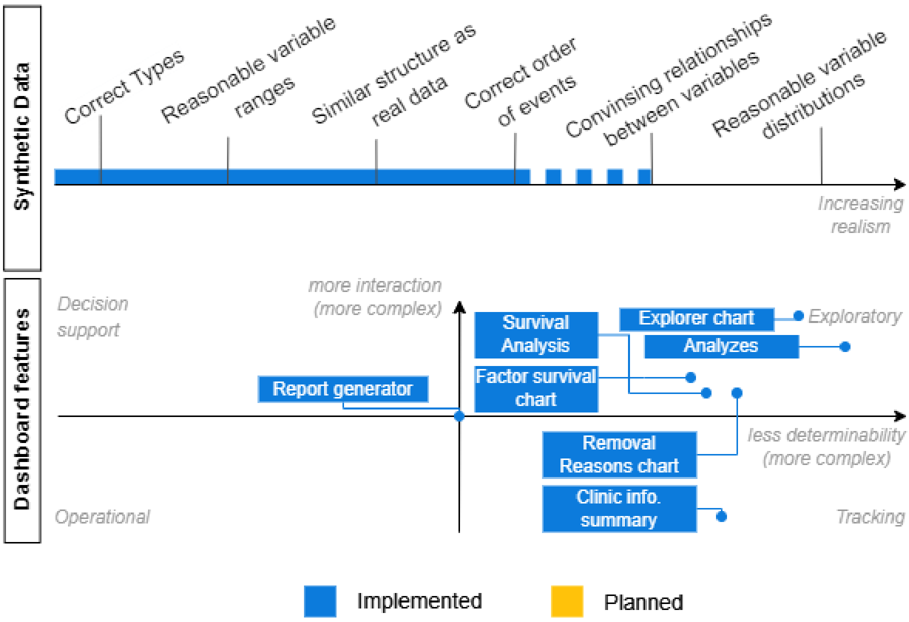


Figure 4.21: Summary of all the implemented features.

their performance. The evaluation revealed that we modified some advanced aspects of the synthetic data. This time, the focus group did not suggest any new features but improvements to the existing features. They said that we should focus less on implant names; instead, we should give more attention to dental implant characteristics. In addition, they mentioned that the charts should compare similar groups of patients and that we should reduce the complexity of some features.

- **Iteration 5** - The synthetic data set remained unchanged. We modified some of the dashboard's features, adding a global filter to make it easier to compare similar patients. In addition, we reduce some of the graphs' complexity. We postponed the evaluation for the final assessment.

Chapter 5

Implementation

While in the previous chapters, we have already presented some aspects of the dashboard application and synthetic data, we will give a more in-depth presentation. We provide a brief overview of the technologies we used and the features that we implemented. Next, we explain the dashboard's architecture and how the dashboard processes requests. Finally, we review the application's extendability, scalability, potential optimisation, and security issues.

5.1 Technologies

We developed both artefacts using *R*, a language and environment for statistical computing and graphics (R Core Team, 2021). The language is flexible and comprehensive, providing tools for data manipulation, calculations, and graphical displays (R Core Team, 2021). *R* has some easy to use and powerful functions to create data summaries. These qualities make it ideal for analysing and visualising data. In addition, it has a considerable amount of extra packages extending the functionality, for example, allowing for the creation of web applications (Jyotsna, 2014).

To generate synthetic data, we mainly used functions from *R*'s standard library. We used a combination of sampling functions and random number generators.

For the dashboard, we used *Shiny* and *ShinyDashboard*, which are frameworks for *R* that make it easy to build interactive and dynamic web applications ("Interact. Analyze. Communicate" n.d.; *shinydashboard* n.d.). These frameworks facilitate interactivity using a reactive programming model with reactive values, expressions, and observers. Therefore, when changes are triggered somewhere in the reactive stream, such as an input value provided by the user, the server may generate a new response (*Reactivity - An overview* 2017). Because *shiny* has implemented this functionality behind the scenes, developers can focus on business logic, mostly ignoring the reasoning of the client-server communication.

To quickly run and host the dashboard on other systems, we used *Docker*. *Docker* packages software into images that contain all dependencies, such as code, libraries and settings, necessary to run the application (*Use containers to*

Build, Share and Run your applications n.d.). Hence, it runs reliably from one environment to another. In our case, the docker image size was relatively large but allowed us to run it on various hosting platforms, such as Microsoft Azure.

5.2 Synthetic data

Although describing the final implementation of the synthetic data set is challenging, we will present some of its key characteristics. Throughout the iterations, its realism has steadily increased due to the attributes that we implemented. After completing the iterations, the artefact has the following properties:

- The synthetic data variables had the correct type, allowing us to insert them into the database.
- The ranges were mostly reasonable, making it difficult to distinguish between real and synthetic values.
- We identified relationships between variables, ensuring that nullable variables were present only if their dependencies were fulfilled. Consequently, the synthetic data mimic the structure of real data.
- The order of events occurred chronologically and had reasonable timespans.
- We implemented a few relationships between variables. We linked implant insertions and removals to patients and added a lower survival probability for some implants. However, we did not model the relationships between the majority of variables due to the complexity of the task.

The variable distributions were unknown because we had no access to a comprehensive real data set. Hence, we did not try to implement this attribute.

5.3 The dashboard

We present the application in this section by first showing the different features before discussing the application's architecture, scalability, adaptability, and security.

5.3.1 Features

Each feature listed below does have a dashboard page. We did not arrange the pages presented in the application in any particular order since the prototype's main objective was to demonstrate functionality. Consequently, any production-ready application will likely use a subset of the features. A common setting across all the features is the global filter. Its main aim is to make it easy to compare similar groups of implants, clinics, and patients.

Explorer chart

- Interaction: High

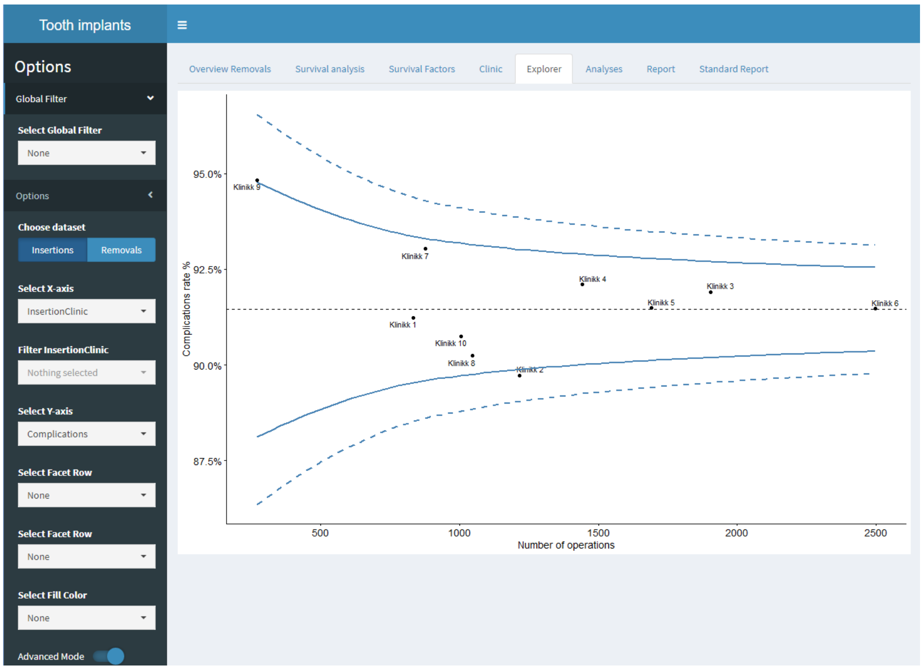


Figure 5.1: The final version of the explorer chart feature

.

- Determinability: Low

The exploratory chart shown in Figure 5.1 allows the user to investigate the data stored in the dental implant registry. Initially, it has few options, where users can choose only one variable to view its distribution and whether they want to view insertions and removals. If the user enables advanced options, the dashboard reveals additional options, giving the user complete control over which variables to use for the x-axis, y-axis, colour, and facets. However, users do not control which graph displays the data. This selection automatically depends on whether the axes are logical, numeric, or factors.

Analyses

- Interaction: High
- Determinability: Low

Figure 5.2 shows the analysis feature after creating a model. This feature gives much freedom to the user, allowing them to specify the method and variables to use in their model. The cognitive load required to interpret the result of the computation may be substantial. When creating a model, the dashboard will show a graph, if possible. In this prototype, we do not support the creation of models with interactions.

Survival Analysis

- Interaction: Medium

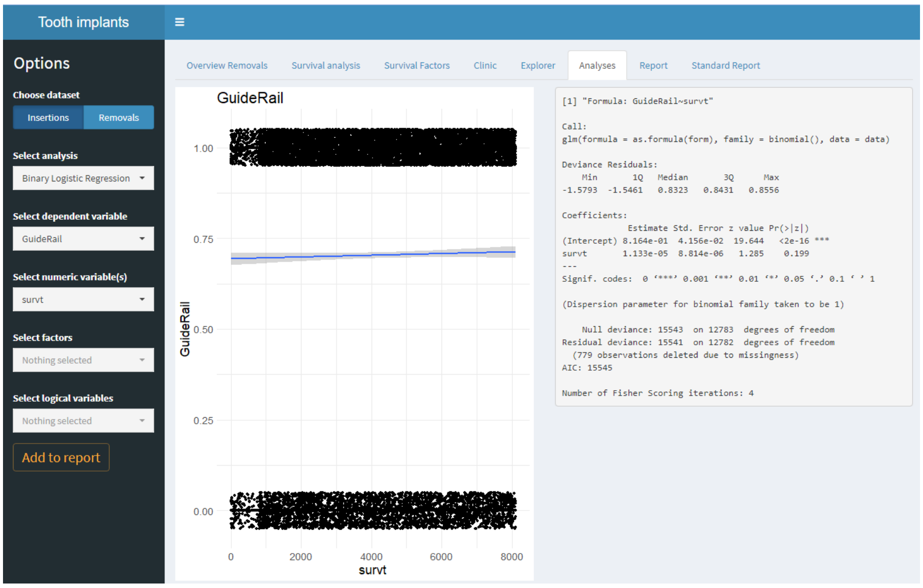


Figure 5.2: The final version of the analysis feature

.

- Determinability: Medium/Low

The survival analysis, shown in Figure 5.3, requires some user interaction and has medium/low determinability. Hence, the cognitive effort to interpret the result may be substantial. The feature computes an estimate of a survival curve based on the selected factors. If no elements are selected, the application calculates the overall survival. Users can also choose the censoring time, that is, the duration of the study period.

Factor Survival

- Interaction: Medium
- Determinability: Medium/Low

This feature compares the survival rate of dental implants after two time periods. Dental implants are grouped according to a factor chosen by the user. In addition, the users are free to select the length of the two time periods, with the default being 2 and 5 years. Moreover, the user can drill down on particular implants (Lot Nr) and view their survival rate with respect to the selected factors. Figure 5.4 shows the survival rate of dental implants grouped by position, with no drill-down option chosen.

Clinic information

- Interaction: Low
- Determinability: Medium

This page, shown in Figure 5.5, contains much information visualised using

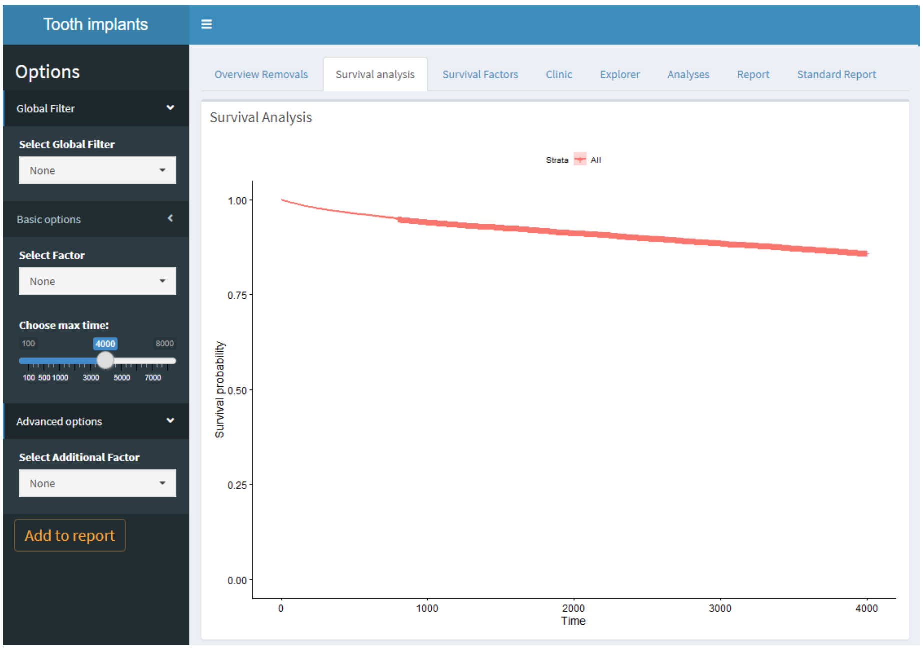


Figure 5.3: The final version of the survival analysis feature

.

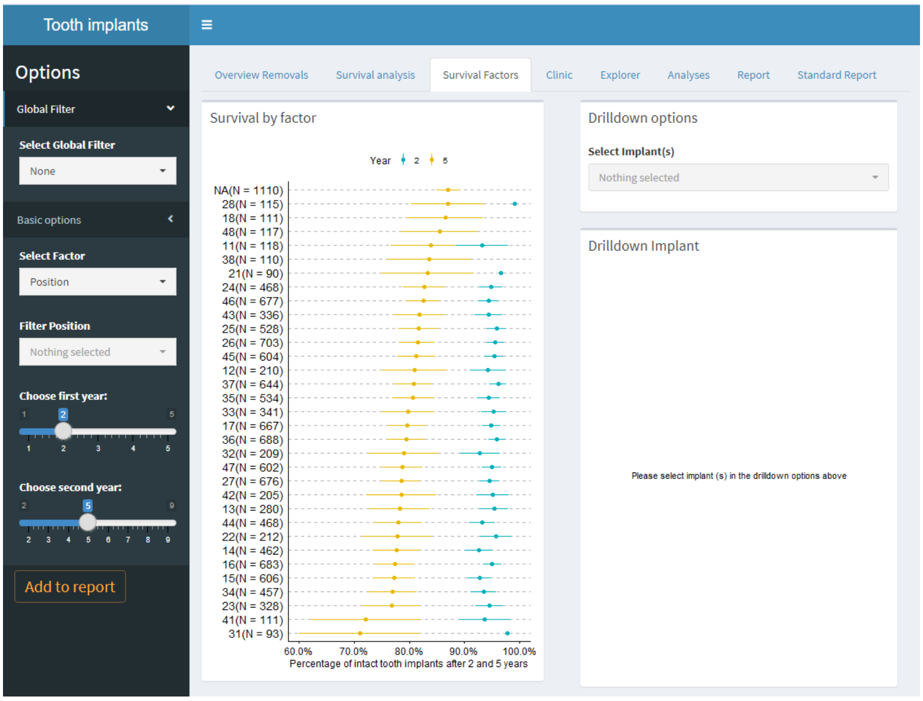


Figure 5.4: The final version of the factor survival feature

.

Figure 5.5: The final version of the clinic information page

.

different components. Potential quality indicators are summarised in infoboxes and coloured according to the result. Green indicates that the specific clinic follows the guidelines, while orange and red indicate a weaker performance. Furthermore, the two following charts in the first row show the last dental implants used by the clinic and the clinic's relative success compared to other clinics. Lastly, the application displays the changes over time at the bottom of the screen. Currently, the application surveys antibiotic usage and first-year removal caused by infections. Users interact only with the feature by selecting which clinic to view and applying the global filter.

Removal Reason chart

- Interaction: Medium
- Determinability: Medium/Low

Figure 5.6 shows the removal reason for dental implants grouped by year. The user can filter the data by removal reasons and year. In the advanced options, it is possible to add a facet row. This option entails exploring how the removal reason changes by year for a particular factor.

Report generator

It is difficult to categorise the custom report generator as it depends on other features. When navigating the application, users can select charts that they want to include in a report. The graphs chosen by the users are summarised on the "Report" page. In this version of the application, it is impossible to edit the selection.

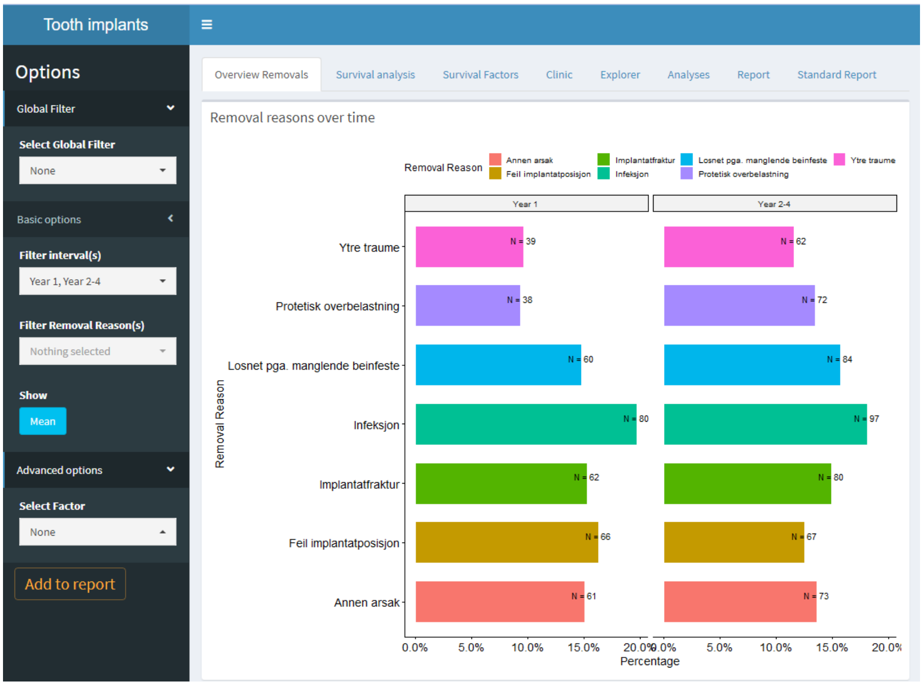


Figure 5.6: The final version of the removal reason chart

.

Standard Report generator

The standard report generator requires little user interaction, only requiring users to select which clinic to customise the report to. Subsequently, the application will generate a selection of pre-determined graphs. These graphs should inform a clinic about its performance and highlight areas of improvement.

5.3.2 Architecture

We provided an overview of the system's architecture in section 3.1. In this section, we have included a flow diagram to illustrate the interactions between the various instances in the architectural diagram (see Figure 5.7). When the user first accesses the web page, starting a session, the server returns the layout, including the predefined UI elements. Following this initial return, the server requests all available data from the database. Additional UI elements are created on the server and sent to the client based on these data. Lastly, the application generates and displays graphs for the client after receiving all required data. A new plot is rendered and returned if the client changes any reactive values by interacting with the UI.

5.3.3 Scalability

Although we developed a dashboard prototype, developers may use part of this application in future implementations. Therefore, it is relevant to discuss the scalability of the application. We can consider scalability regarding the number

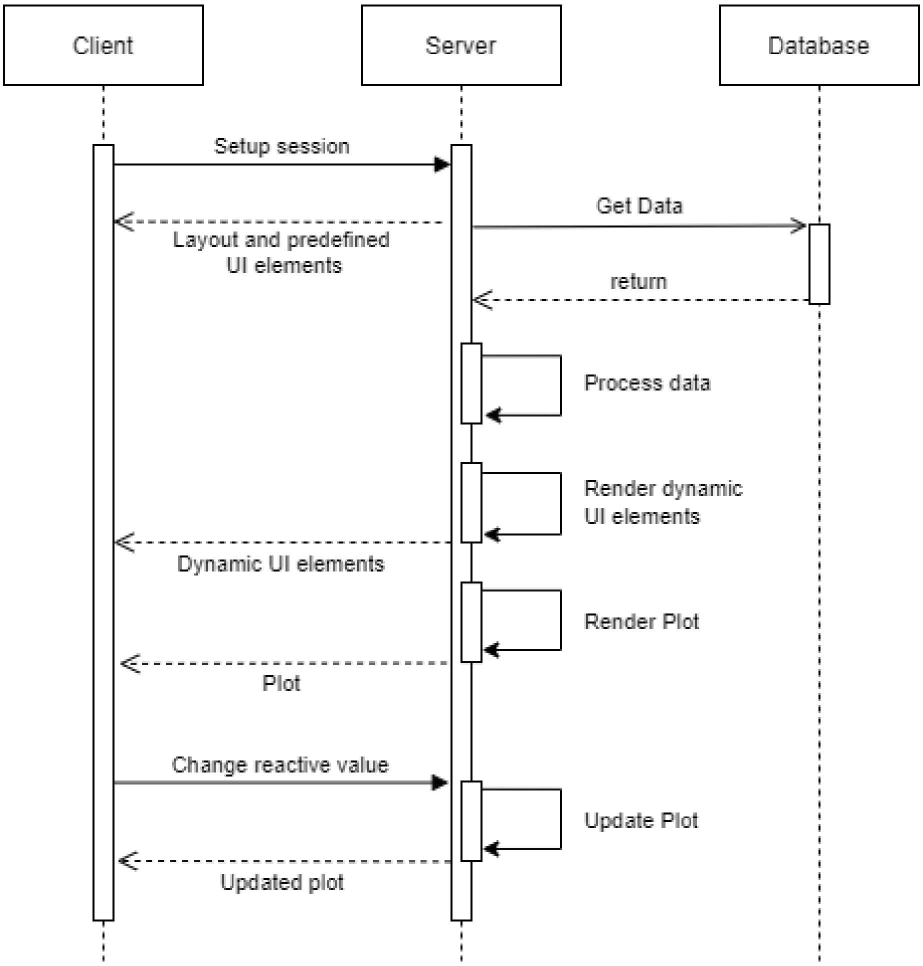


Figure 5.7: A flow diagram of a user initiating a session and interacting with a plot.

of entries in the database and simultaneous users.

At the start of a user session, the application makes a single call to the database to transfer all data related to implants, insertions, and removals stored in the registry. Therefore, the server transmits a lot of data during this process. This transfer enables the exploratory features of our dashboard, as users can define their visualisations. However, it may affect the application's scalability. We measured the ability to scale in the final evaluation in Figure 6.5.

Our artefact handles user requests sequentially because R is single-threaded (RStudio Support, 2022). Potentially, users had to wait for other users' computations to finish before their own could start, slowing the application. However, there are still benefits of using a single-threaded approach, such as taking less time to start, using less memory, and avoiding re-running part of the code (RStudio Support, 2022).

Generally, the artefact's performance was satisfactory with few concurrent users, even for larger datasets. Much performance can be gained from a hosting platform, which can potentially run several images in parallel. Consequently, we allocated little time to optimisation. However, developers might need to take measures to reduce response time as the registry grows. Below are some suggestions.

Developers may consider application-level cache to avoid re-fetching the data for each user as a first potential optimisation measure. The server saves the registry data in memory or a local file using this method. It is important to note that, when using memory, the cache does not persist if the R process stops, usually when no users are connected. After an appropriate amount of time, the application invalidates the cache, and the first users have to send a request to the database again. However, as the database grows large, using a memory cache to store data is infeasible due to its small size. Still, we may use the memory cache to store plots or other small items. A local file cache may be a viable option for larger databases. This alternative has the additional benefit of persisting after the R process stops, potentially saving us from the need to fetch data from the database. In the case of a large database, developers may implement other measures to ensure optimisation, for example, using more fine-grained SQL statements.

Another potential optimisation is to debounce user input. A debounce is a higher-order function that returns another function after a given time interval. The returned function is the reactive expression containing the user input in our case. However, every time the user calls the debounce function within the specified time interval by changing the state, it postpones the return. Therefore, the frequently changed input will not immediately trigger new data calculations. This delay saves resources by delaying performing an expensive computation until the user finishes editing, but only if the user updates within the debounce interval. An alternative is to let the user click on a button after they have finished setting all parameters.

5.3.4 Adaptability

Even though we developed the application for a dental implant registry, some functions are easily adaptable to similar registries. Due to design choices, this flexibility is particularly true for the more general parts of the application, such as the "Explorer" feature. The options given to the user are not hardcoded but based on the data table's definitions. Consequently, the possibilities are derived from the input data, needing some load time before being outputted to the user. The y-axis, x-axis and other dimensions of the exploratory graph require only these input variables and the main data. Therefore, the implementation is agnostic to the data, and any registry could use it.

5.4 Security

We label any security measures outside the scope of this project. However, because the dashboard will have to handle actual patient data in the future if utilised, we will discuss some security and privacy issues.

The features we implemented have to require authorisation. Specifying privileges is necessary to ensure that users only access the resources they are allowed to view. Permission to use all features should be given to registry administrators while being limited to other stakeholders. For example, clinicians should only be allowed to view their clinic's performance compared to the mean, not compared to other clinics. Consequently, developers must identify the appropriate access for each stakeholder. In some cases, however, it might be sufficient to reduce the freedom given to specific stakeholders by removing certain options.

Although the application displays primarily aggregated data, there is a slight possibility that patients may be identified. As a rule of thumb, aggregated data with less than ten observations should be excluded from charts when this possibility exists and the user does not have proper authorisation.

Ideally, the dashboard will be easily accessible to stakeholders in their daily workflow. This incorporation could entail using the same authentication method as their other systems.

Chapter 6

Final evaluation

Evaluation occurred as an intricate part of our iterations; additionally, we concluded with a comprehensive evaluation of our two artefacts at the end of the project. Mainly, we assessed whether our solution satisfied the objectives of the solution mentioned in section 1.3.2. First, stakeholders reviewed the implementation of both artefacts through semi-structured interviews. Furthermore, we tested the realism of synthetic data using a Turing test and evaluated the user experience of the dashboard using System Usability Scale (SUS). Lastly, we evaluate the technical performance of the dashboard under various conditions.

6.1 System implementation

A highly relevant assessment of our system was whether the implementation of our two artefacts was valuable to identify potential useful dashboard features and determine whether synthetic data supported the development of the dashboard. Consequently, indicate whether we achieved objectives 1 and 2.

We conducted five semi-structured interviews to investigate these aspects. This format allows for specific and open questions, bringing forth foreseen and unexpected types of information (Hove and Anda, 2005). The five participants were the stakeholders involved in the development process (listed in Section 3.2).

We conducted the interviews individually and recorded the interaction. The interviews were structured as follows:

1. A short presentation of the artefact, where we showed the menu and navigation options.
2. The participant tested each feature of the dashboard. During the test, we used the think-aloud method (Lewis and Rieman, 1993). This method required them to explain what they were trying to do and encouraged them to ask questions that arose during the testing.
3. After testing the application, we asked the questions presented in Appendix A

We translated and summarised the responses from the think-alouds and interviews. In Table 6.1, we recap the responses of the participants regarding the usefulness of the dashboard features. Any suggestions made by the participants during the interview are included in Table E.1. Finally, the comments of the respondents related to the synthetic data are listed in Table 6.2.

6.2 Turing test

To evaluate the realism of the synthetic data, we performed a Turing test with actual and synthetic data. This approach has been proposed to compare these types of data (Chen et al., 2021).

Real data was collected as part of the initiative to create the dental implant registry. It was collected using a questionnaire with open-ended and closed-ended questions. Adapting the data to the registry was a real challenge due to open-ended questions. We had to spend a lot of time cleaning the answers. For example, the respondents had written the name of a clinic in more than ten different ways. Consequently, we had to merge these variants into one to ensure the validity of the data. After we had prepared the data, they contained around 1200 observations, most of them being insertions.

We structured the trial as follows:

1. A domain expert first used the "Removal Overview" feature in two sessions for a short period. In one of the sessions, the data was real, while in the other, it was synthetic.
2. After completing the two sessions, the expert had to guess which session he interacted with real and synthetic data and then give a brief comment on his selection.

We used two domain experts; one of the researchers and the clinician. After reasoning about the distributions visualised in "Overview Removal", these experts identified the real and synthetic data. They used some time to reason and think before announcing their answer. A suspiciously large percentage of certain removal reasons made it possible to differentiate.

6.3 Stakeholder mapping

We wanted to create an overview of the features that may be suitable for particular stakeholders. Therefore, after testing the application, the various stakeholders completed the table in Appendix B. Beforehand, we selected five potential stakeholders. However, the interviewees could also mention others.

We show the results in Figure 6.1. One of the statisticians selected "others", arguing that journalists and media might be potential stakeholders interested in the features.

Table 6.1: Summary of interviews and think-alouds about the usefulness of the application features

Useful	ID	Synthesis
Overview Removals		
Yes	S2	This feature and similar features which provide an overview are valuable.
	R2	It may increase the understanding of what and why implants was removed
Survival Analysis		
Yes	R2, C1	May benefit researchers with knowledge of statistics
	S2, S1	It is interesting and valuable to create such figures to explore the data quickly.
Factor Survival		
Yes	S2	Informative feature, making it easy to compare survival rate of different factors
	R1	The functionality offered by this feature seems valuable
	S1, C1	It seems like a clear and informative feature
Clinic Information		
Yes	S1, C1	It allows clinics to assess their performance, and compare it against others
	S2	It is informative and easy to understand
Analyses		
No	S1	Unable to use generated models in research as this requires more processing of the data.
	S2, R2	Too technical, requiring special skills to interpret
	R2	Difficult to adapt to your needs.
General observations		
Yes	R1	Features that can be used to identify characteristics with a higher probability of fault
	C1	Features that overview patients and the performance of the clinician. If there is significant deviations in performance, the dashboard should have features to determine the cause.
No	S2	More technical features that require special skills to interpret the results are less useful.
	C1	Difficult to identify non useful features because much depends on area of use.

Table 6.2: Answers concerning the synthetic data

ID	Synthesis
Bothered by synthetic data?	
S1	Impossible to know whether the data were real or false without domain knowledge.
S2	Synthetic data help evaluate features because the focus remains on features, not data.
R1	No, difficult to judge the realism of the data.
R2	Noticed some irregularities, such as many long implants, but was not distracted by them.
C1	No, I did not study the data, only the features were in focus during this session.
Synthetic data useful through the iterations?	
S1	Useful to demonstrate the features.
S2	Maybe even more useful than real data because the focus remains on the features and not on interpreting the data.
R1	Exciting to watch demonstrations even though the data was synthetic.
R2	Yes, it would have been better with real data, but lacking such data the synthetic data proved helpful.
Synthetic data's realism affects its usefulness?	
S1	Too little domain knowledge to differentiate.
S2	Demonstrating some features, such as survival analysis, would have been less interesting if we had used random data.
R1	Completely random data would have been difficult to relate to.
R2	The synthetic data used in this project was an improvement compared to random data, allowing us to accomplish more.
C1	The realism of the synthetic data did not affect him in the evaluation. However, he adds that it might have been distracting if we had not informed him beforehand.

Figure 6.1: A heatmap of which features the respondents deemed suitable for various stakeholders.

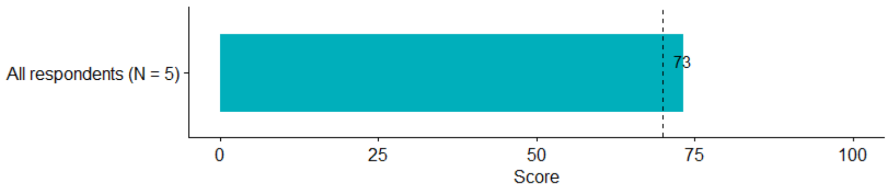


Figure 6.2: The mean System Usability Scale score for all our respondents. Scores above the dotted line indicate acceptable performance.

6.4 User Experience

We quantified the user experience using SUS, a tool developed by John Brooke (1996). According to Bangor et al. (2008), a score greater than 70 indicates that the performance of the tested application is acceptable. A rating above 90 indicates superior products, while a rating below 50 indicates unacceptable products. We show the application's SUS in Figure 6.2.

6.5 Performance

We present the performance of the dashboard through a load test. This test helps developers estimate the number of users that their service can support. In addition, it can identify the main performance bottlenecks and guide improvements to the artefact.

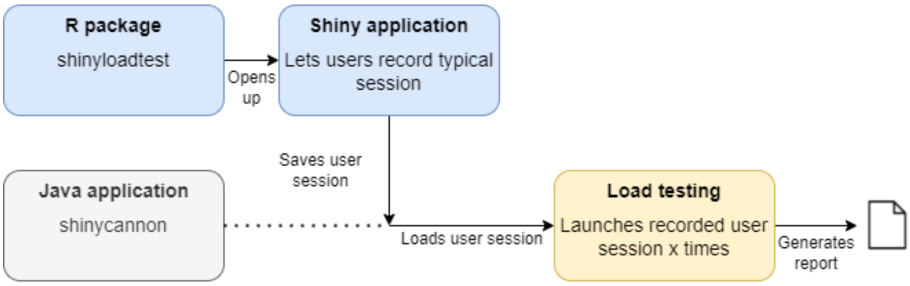


Figure 6.3: Load test setup

The load test was performed on our local machine running a docker container with the application. This setup ensures minimal latency in a client's communication with the application. Hence, we reduced the factor of a server's varying response times and instead tested the artefact's time to process the user's request and respond. The dashboard communicated with an SQL database hosted on Microsoft Azure, which contained 1500, 15000, or 75000 dental implant insertions and removals.

Figure 6.3 shows the test setup. We first recorded a possible interaction between a user and the application using *shinyloadtest*. In this recording, we interact with various plots by adjusting their parameters. Subsequently, we used *shinycannon* to run this recording and emulate 1, 4, and 16 users.

The results shown in Figure 6.5 indicate that the application can handle four users with a minimal increase in the time compared to one user session. However, with 16 concurrent users, the time of a user session increases dramatically. The original user session that lasted 100 seconds now takes three times longer before being completed. The main bottleneck of the application appears to be the first load when creating a user session (Figure 6.4). At this point, the application retrieves all the necessary data from the SQL database and prepares them for use. Consequently, a lot of computing power is initially required, as is evident in Figure 6.4 and in Figure 6.5 at 0 seconds.

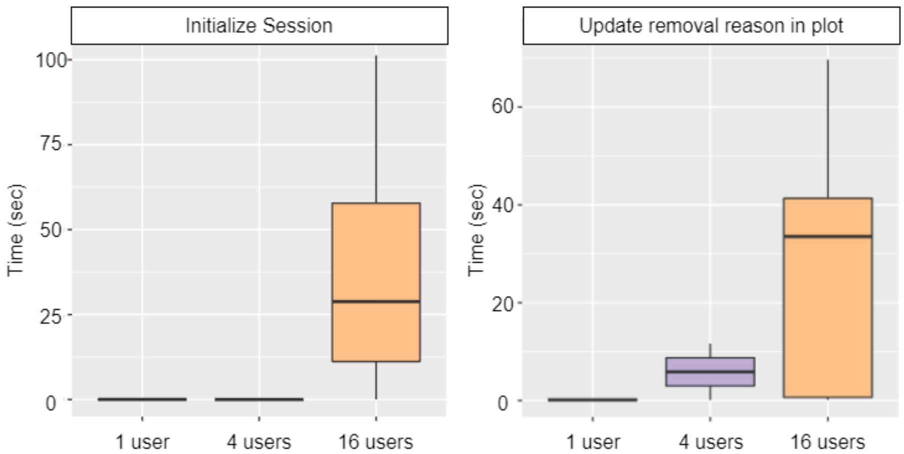


Figure 6.4: We selected two events from the load test with 15000 database rows to show the variability in their completion time.

Figure 6.5: We populated the database with 1500, 15000 and 75000 rows of insertions and removals, and then we performed load tests simulating 1, 4, and 16 concurrent users. A red line shows the running time of the original script. Sessions exceeding this line suggest an overloaded server.

Chapter 7

Discussion

In this chapter, we will discuss the findings discovered through this project. We will base our discussion on how the artefacts solved the objectives listed in Section 1.3.2. Mainly, we will focus on key results and their implications. At the end of the chapter, we list the limitations of the project and give a brief recap of our experiences using design science.

7.1 Synthetic data and development

The project's objective related to the synthetic data was to determine an approach to generate synthetic data that can be used in a functional prototype, supporting the development of registry services when real data are unavailable.

We described an approach for developing secondary services for a dental implant registry using synthetic data generated manually and a high-fidelity prototype. Together, the synthetic data and the dashboard prototype can engage stakeholders early with a working application rather than just conceptual ideas, which may be valuable for novel services. In addition, using evolutionary prototyping, developers may use parts of the artefact in a future application.

The results indicate that this approach can help to discover potentially useful features and functionality and, furthermore, inform potential users about such features and functionalities, and then together expand further on the development of future services. We were able to generate a synthetic data set that supported this process. To overcome the challenge of manually developing synthetic data, we used an iterative approach that increased its level of realism when needed. This method limited the amount of work we used to create the data and provided us with insight into how qualities related to the synthetic data affected the development.

The development process

Usually, part of the development process consists in obtaining requirements that describe what a system should do (Bourque et al., 2014). In this process, stakeholders contribute to the application requirements with their different points of

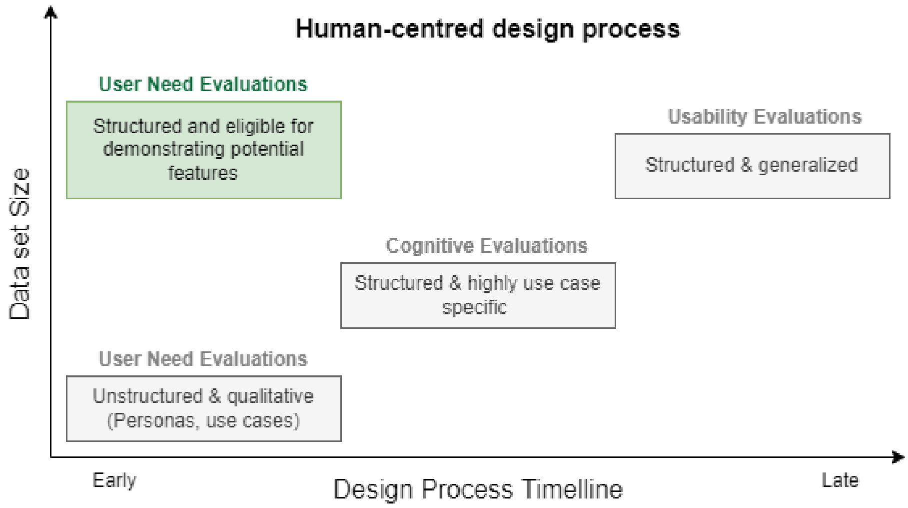


Figure 7.1: The data requirements needed to perform user evaluations changes throughout the design process. Typically, the first stages need less data than the later stages in the human-centred design process. However, the green box highlights our approach, using a large synthetic data set to demonstrate potential features early. This approach was feasible because the application's novelty could have made it challenging engaging users with conceptual ideas.

view (Bourque et al., 2014). For example, the early stages of a human-centred design process focus on qualitative and unstructured data, making personas and developing use cases, thus understanding user needs. Subsequently, developers can create prototypes to test design ideas (Preece et al., 2015). However, this approach was less feasible in our project because of the system's novelty, with users unsure of the possibilities related to the registry's secondary services.

To overcome these challenges and uncover the requirements for such services, we instead developed a synthetic data set and a prototype to be used in early demonstrations and evaluations. Kinnaird and Romero (2010) recommend using a prototype as soon as it exists to understand the interest of users. We take this further in our approach, suggesting that the prototype should be used at the beginning of the design process to understand users' needs and the application requirements. Figure 7.1 is adapted from Pollack et al. (2019) and shows the data requirements through different phases of a human-centred design process. The green box highlights our approach in this context.

We used a relatively sizeable synthetic data set and a functional prototype. Consequently, we were able to simulate years of patient data recorded by the registry and use these to demonstrate possible secondary services. This approach was particularly valuable when the prototype was an innovative and unique dashboard within implant dentistry with several exploratory and analytical features. It allowed us to truly show these aspects of the dashboard, enabling data exploration during the demonstrations and the evaluations. The results indicate that this functionality motivated and engaged our stakeholders, resulting in them suggesting new features, functionality, or improvements to the prototype.

These findings support the idea that stakeholder participation is a critical factor in successful software design (The Standish Group, 1994). Moreover, it is consistent with the observation of Kinnaird and Romero (2010) that it is possible to gain insight into users' interests by combining a focus group and prototyping activity. Furthermore, we observed some of the same benefits of prototyping as Nelson et al. (2016), even though our project used synthetic data. First, some of our observations support the idea that early prototyping motivates users to contribute ideas for improvement. For example, in iteration 3, one of the statisticians suggested a survival analysis feature after we demonstrated a similar feature (Table 4.6). Second, it was easy to identify shortcomings in the prototype during the evaluations, as is evident from the list of improvements and suggestions made by stakeholders in Appendix E.1.

Generating Synthetic data

Creating an early prototype can be challenging for some projects, making our approach less feasible. We suggest developing the two artefacts in parallel to try and overcome this challenge, ensuring that the synthetic data set and a prototype are ready simultaneously.

It is crucial to streamline producing a synthetic data set to reduce delays in prototype development. We outline such a process throughout the iterations. We found that generating synthetic data with *R* was the most straightforward. However, there are applications dedicated to manually generating synthetic data, such as those proposed by Mannino and Abouzied (2019) and Althar and Samanta (2021). The benefit of using these types of applications is that developers are likely to get a better overview of the variables and their relationships. Additionally, it may be easier to reproduce the data by communicating the settings used to generate the data set. The downside is lower flexibility.

Flexibility was the most influential when we decided to use *R*. We could easily use and combine functions from the standard library and external packages to get the desired results. Moreover, it was useful to freely iterate through and manipulate the data set when introducing interactions. Also, since visualising the data is easy with *R*, getting an overview of the created data is possible.

Some studies discuss synthetic data generated manually in software development, but none has used them similarly to this project. Instead, developers often employ these data to create specific scenarios to evaluate the suitability of an application to solve various tasks (Whiting et al., 2008; Pollack et al., 2019). Although the main focus of our approach is to identify the requirements of different services, developers may later use it in scenario testing to evaluate multiple designs. We introduced some relationships in the data during iteration 4 by reducing the survival rate of various implants to test the survival analysis feature (Table 4.8), indicating the potential of generating data for scenario testing. However, compared to Pollack et al. (2019), the scenarios are likely to be less detailed due to the amount of data we create.

Level of realism

The level of realism is a fundamental aspect likely closely related to how synthetic data can support the development of registry services. First, the degree

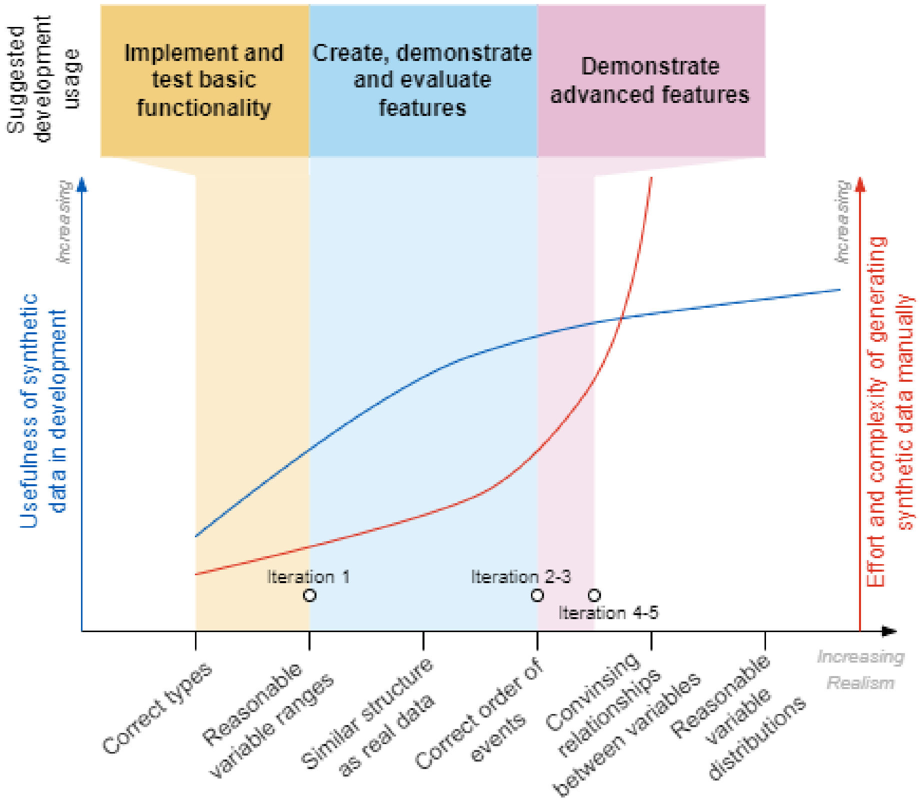


Figure 7.2: Shows how the realism of the synthetic data affects its usefulness in development and the effort of producing the synthetic data. The highlighted areas indicate potential development usage for synthetic data with different attributes. The points mark how the synthetic data progressed throughout the project.

of realism may affect the usefulness of the synthetic data in this process. By usefulness, we mean how synthetic data contribute to engage stakeholders and influences the quality of their feedback. Second, more realistic data are increasingly difficult to create, requiring developers to carefully consider their level of investment in producing such data. We summarised some of our findings and observations on these factors in Figure 7.2. It is important to note that the trends shown in the graph are uncertain, as we do not have a precise measurement of effort and usefulness. Nevertheless, the chart indicates the overall pattern observed in this project.

In cases where the synthetic data have a low degree of realism, we found it useful to build fundamental aspects of the application, such as creating and testing calls to the database. This type of data requires minimal effort to produce. However, it can distract stakeholders if used to demonstrate and evaluate features, resulting in less predictable but potentially valuable feedback. We made this observation in iteration 1, where there was a notable contrast between the complexity of the two artefacts that we developed. The synthetic data were simple with a low degree of realism but had correct types, and most variable ranges were reasonable. In contrast, the dashboard was already quite complex, allowing users to explore the data freely. Consequently, combining these artefacts probably highlighted irregularities in the synthetic data. These inconsistencies may have resulted in the feedback group shifting its focus from the exploratory graph to the synthetic data during an evaluation meeting. Hence, the participants' responses did not directly discuss the intended feature. However, the feedback was still valuable, alerting us to additional options necessary to properly explore the data.

After modifying the synthetic data to have a structure more similar to real data, the group focused on the intended features in the rest of the evaluations. Hence, the usefulness of the synthetic data set increased, as shown in Figure 7.2. Of course, this may also result from the focus group getting more used to the evaluation format and ignoring any irregularities in the data. A more plausible explanation, supported by the interviews, was that the realistic qualities of the synthetic data were adequate to evaluate and examine the dashboard features. Developers can quickly produce this type of data. The implication is that developers should aim for some structure when generating synthetic data to create, demonstrate, and evaluate features. However, data with a higher or lower degree of realism are still valuable.

Although we further improved the realism of the synthetic data set in later iterations, it was difficult to determine how this affected the evaluations. There did not seem to be any detectable difference in the feedback from the focus group after the synthetic data reached a certain level. However, improving the realism of the data helped to adequately demonstrate more advanced dashboard features. For example, one of the statisticians commented that the survival analysis demonstration was more interesting because the shape of the curves was recognisable. The downside of producing more realistic synthetic data is that the effort and complexity of this task are massive. Figure 7.2 reflects these observations. Consequently, developers must consider the cost and benefits of producing such a synthetic data set to be used in development.

From the interviews, there was a less clear trend related to the usefulness of synthetic data in development, raising some doubts about the trends we observed and summarised in Figure 7.2. In general, stakeholders perceived the synthetic data as valuable throughout the development process (Table 6.2). Interestingly, one of the statisticians said that he preferred synthetic data over real data, as this ensured that the focus remained on the features during demonstrations. One of the researchers would rather have actual data, while the clinician was indifferent to the data used in the evaluation. The clinician's perspective differs from that of the other stakeholders. This contrasting perspective is also evident when discussing data without any realistic qualities. Although most stakeholders agreed that such data could potentially interfere with the development process, the clinician did not share this perspective. However, he interacted with the dashboard after the synthetic data set had reached a quite high degree of realism, which may have affected his assessment.

Consequently, more research is needed to verify the trends we observed through the iterations. This research includes proving whether there is an advantage in generating synthetic data with some realism and whether the level of realism of the synthetic data may support different development phases. When researching this, it is essential to note that the trends that we summarised in Figure 7.2 may differ in other projects. This difference may be due to the variation in the system complexity and the knowledge of the stakeholders involved.

Section Summary

The key takeaway is that the development of a synthetic data set and its use in early demonstrations may generate valuable responses from stakeholders, revealing the needs of potential users in a novel application. It may be beneficial for the data to have enough realism to demonstrate the features adequately, ensuring more precise user feedback. However, inconsistencies in the synthetic data are not necessarily damaging, as they may initiate a discussion and highlight other aspects of the application that need consideration. Nevertheless, developers should use caution before investing too much time in synthetic data, as the trends observed in this research need to be verified, and the value of synthetic data may vary from project to project.

7.2 The dashboard

The dashboard's objective was to identify potential features and functionality that can be helpful to stakeholders in exploring data from a dental implant registry, laying the foundation for future registry services.

We developed six features that we classified according to the categories presented in Zhuang et al. (2020). Additionally, we created two report generators that this template could not classify because they display data generated by the other features. The results indicate that most of the features we developed were perceived as useful by the stakeholders (Table 6.1). There was a tendency for stakeholders to consider functionality with a low degree of complexity as the most valuable, while more advanced features received little or negative feedback. The following is a summary of the features we developed.

- **Explorer** - Freely explore the data in the quality register.
- **Analyses** - Create statistical models and test significance.
- **Survival Analysis** - Investigate factors influencing the time from insertion to the removal of a dental implant.
- **Factor Survival** - Compare the survival rate of dental implants after a given time, grouped by factor.
- **Removal Overview** - Distributions of removal reasons for dental implants grouped by years since insertion
- **Clinic Information** - Summarises statistics for each clinic, reporting information about their performance
- **Report** - A customisable report created by the user.
- **Standard Report** - A fixed report with predefined graphs and information.

By examining some of the key motivations for a dental implant registry, we can probably explain why many features were considered beneficial. A dental implant registry should help identify underperforming implants and unfavourable characteristics of implants. In addition, it should allow clinics to assess and improve their performance. The functionality offered by "Overview Removals", "Survival Analysis", "Factor Survival", and "Clinic Information" all supports these goals. Consequently, with these features adequately implemented (SUS score in Figure 6.2), stakeholders can use the application to investigate some of the fundamental aspects of the domain.

Although most of the features we developed are not directly comparable to other studies, 'Clinic Information' has parallels to the dashboard created by Weiss et al. (2018). This study found that the dashboard positively affects the quality of care. The dashboard reported quality indicators using colour to grade the unit's performance, similar to ours. However, the difference is that we supplemented the dashboard page with additional functionality. For example, we added a funnel graph to compare the survival rate of dental implants inserted by a clinic with other clinics. In addition, we used a chart to show changes over time in the use of antibiotics and first-year infections. Several stakeholders highlighted this functionality as valuable. These responses are in line with a suggestion by Benoit et al. (2022) that practitioners should have access to indicators that allow them to improve. Consequently, the 'Clinic Information' and similar functionality could help increase the quality of care at different clinics, and developers should consider multiple such data visualisations when communicating clinics' performance.

Complexity

The results show mixed responses when comparing the simplest and most advanced features. Although the most straightforward feature, the "Clinic Information", received positive responses, the more advanced, the "Analyses" feature, was the least helpful (Table 6.1). The complexity of each feature is probably an important factor in explaining this trend. To some extent, stakeholders perceived the more complex features as less valuable. To use the "Analyse" feature, a stakeholder requires knowledge about the domain and statistics to analyse the results. Furthermore, any created models must be interpreted with caution and should only be used to identify areas of interest. The challenge of using such a feature appropriately is not easy to overcome.

Consequently, this result indicates that exploratory features with statistical tests should not be a priority when developing services for a dental implant registry, at least before there is a clear need. These features are usable only for a few highly competent stakeholders and may require advanced options to fine-tune the analysis. However, it is unlikely that even stakeholders with the necessary proficiency can use it to the full extent due to the data processing required to create publication-ready models and graphs. Hence, the application's analytical features fall between stools, neither advanced nor straightforward enough.

However, developers should not necessarily dismiss all statistical tests. The respondents appreciated the "Survival Analysis" more than the "Analyses". This feature may be easier to understand and is always accompanied by a graph.

Therefore, it may be helpful for more stakeholders as long as they are aware of its limitations. An area of use is to quickly explore the data. Although developers may include such features in the application, they should also evaluate simpler alternatives. One of the statisticians noted that the "Factor Survival" feature might replace the "Survival Analysis" because it is more straightforward.

Complexity is only one dimension that can influence the usefulness of a feature. This trend is evident by reviewing the feedback of the "Explorer", which has some of the most advanced functionality of the dashboard's features. Users perceived it as interesting, but neither thought it was particularly valuable or worthless. Therefore, complexity can indicate which services are worth developing but should not be the deciding factor. Proper training and other functions that reduce cognitive load can mitigate the negative attributes of such features.

Interestingly, all the features that were well received had roughly the same level of outcome determinability (Figure 3.3). This perspective might indicate that having some limitations on the graphs might be beneficial because their goal and use-cases are more clearly defined. Another explanation might be that users did not receive enough training to fully understand the potential of the most complex features, which reduced their scores. Consequently, developers should consider the needs of the stakeholders, the complexity of features, and the possibility of training when creating services for a dental implant registry.

Stakeholders

According to our respondents, the "Clinic Information" and "Removal Overview" features were suitable for the widest variety of stakeholders (Figure 6.1). There was a consensus among the respondents that researchers, clinicians, and registry administrators would probably benefit the most from the functionality, getting an excellent overview of the data. Most of our respondents thought that other stakeholders with less domain knowledge would likely struggle to understand these features, even with the features' lower complexity compared to others.

However, one of the statisticians suggested that the patients were potential users, arguing that some of them may want to immerse themselves in this field, despite its complexity. The clinician also mentioned that patients could use the "Clinic Information" feature. He reasoned that patients might want to identify the best clinic. Consequently, tracking-type features or features that provide an overview should maybe be a priority if developers want to target a broader range of stakeholders.

From the stakeholder mapping in Figure 6.1, it is noticeable that professionals are considered the most suitable to use the exploratory features we developed, especially researchers. We expected this grouping because it required domain knowledge and statistics to analyse some of the results generated by the application. In the interview with the clinician, he says that only a few clinicians would probably use these features to obtain information about dental implants in general.

We developed various features that were possible to adapt to several stakeholders. Any stakeholder accessing the prototype has access to all these features. However, developers should select only a subset of the functionality that fits the target audience in a production-ready application. In these cases, it might be beneficial to further adapt the features to specific stakeholders by considering which options to include and how to integrate them into the workflow. Similarly, any features selected for a public website should be adapted to a broader audience. In addition, developers should consider features explicitly designed for patients and the integration of PROM into an application.

Report features

The reporting features consisted of a dynamic report and a standard report. We did not discuss these features in great detail during our final evaluation. Therefore, the feedback was limited. Figure 6.1 indicates that clinicians, registry administrators, and managers are suitable for this feature. Registry administrators are likely to create the report, while other stakeholders may benefit from reading it.

However, the results indicate that the current implementation of our report generator is cumbersome and slow, requiring stakeholders to navigate to each dashboard page and click "Add to report". Therefore, one of the statisticians suggested adding it as a separate feature. This dashboard page might include an overview of the graphs, allowing users to select which to include in a report. Moreover, it should save the previous compositions and settings.

Usability and Performance

Although the main focus was on functionality, stakeholders perceived the features through the UI. Therefore, it is also essential to mention these aspects. The consensus among our respondents was that the layout and functionality were overall good. This result is reflected in the System Usability Scale (SUS) in Figure 6.2. However, there was some room for improvement, as shown in Table E.1.

7.3 Parallel development

The last objective was: Evaluate how the parallel development of the dashboard and synthetic data may contribute to the registry's data model and further develop and verify it.

In general, when creating the dashboard and a synthetic data set in parallel, we better understood the data model, as we learnt how the variables depended on each other. By doing this in parallel, we were constantly required to relate the features to the data model and vice versa. This approach ensured a close connection between these components, making it easier for us to know which variables to process and how to process them in the application. The iterative aspect of this method was also important. When the level of realism of the synthetic data gradually increased, we found it easier to identify substandard qualities related to the data model because we were learning more about it.

The most important result of this process occurred in the second iteration. After receiving feedback in iteration one that the synthetic data set contained some inaccuracies (Table 4.2), we wanted to structure it more like real data by including dependencies between variables (Table 4.3). When updating the data, we identified suboptimal groupings in the data model, which caused many null values in the database. Although developers might have detected these flaws in the data model using other measures, recognising them came naturally when generating the synthetic data due to the need to map dependencies between variables.

Later, we did not identify new optimisations for the data model when creating the most complex synthetic data set with the most realistic attributes. This observation suggests that, after gaining a basic understanding of most aspects of the data model, working more with it may not necessarily help further improve it. It should be noted that these improvements and adjustments may not have occurred if we had made the data model ourselves.

Although we detected potential changes to the registry's data model, none of the focus group participants suggested changes or additions. This result was surprising, as we suspected that the interaction with the data through the application's features might contribute to the data model. An explanation might be that we gave respondents too little time to evaluate and test the rather complex application. If it had been simpler, the time might have been sufficient. More research is needed to determine whether stakeholders can contribute to a data model using a parallel development approach.

7.4 Limitations

The study has several limitations.

The most significant limitation was probably the time constraints. There was limited time to discuss and test the features during the iterations' evaluations. Therefore, participants may not have sufficiently understood all the dashboard functionality. This potential lack of knowledge may have affected their feedback and perception, influencing the results related to the development process and the dashboard functionality.

Furthermore, time constraints may have influenced our ability to extract relevant information from stakeholders. Consequently, we may have created less relevant dashboard functionality, and our observations related to the synthetic data may have been flawed. For example, we did not discuss how stakeholders perceived the synthetic data after each iteration.

In addition, time constraints may have influenced the final evaluation. There was limited time to let the participants thoroughly test the features during this process. The results of the following interviews, SUS, and the mapping of stakeholders could have been different if the participants had been able to study the features in more detail.

Another limitation of our study was the focus group composition. We had some stakeholders with limited domain knowledge and the group consisted mainly of researchers and statisticians. Therefore, the features that we identified as useful could change if the focus group had a different composition of stakeholders or if all were experts in the field. Moreover, the group's composition may affect the value of the synthetic data in the development process.

There are also limitations to generalising the results. We based our research on a data model developed by previous students. This foundation could influence the results of our study, and it is challenging to know whether a parallel development approach is adaptable to other instances where there is no data model in advance. Furthermore, due to the varying complexity between domains and the competence of stakeholders, the trends observed in this study may be different in other projects.

7.5 Use of Design Science

The use of design science helped us in structuring the project. It clearly outlined the cycles needed to develop the artefacts and find solutions to the objectives. This information was important when designing our development process, encouraging us to include aspects such as the knowledge base and internal evaluations. In addition, the iterative aspect of the design science was highly useful for our project, allowing us to gradually improve the artefacts and examine how modification to these artefacts influenced the development process.

Chapter 8

Conclusion

In this project, we explored the possibility of using synthetic data to create secondary services for a potential dental quality register. We had three research questions that we tried to answer.

How can synthetic data generated manually support the development and the demonstration of new services for a potential dental implant registry?

The result suggests that manually generating a synthetic data set may support the development of new services by allowing early demonstration and evaluation of features. Through this process, potential stakeholders can view and interact with a prototype to help uncover possible application requirements. However, we found that a synthetic data set should have some realistic qualities to best support this approach, ensuring relevant stakeholder feedback. These qualities include variables with the correct types, reasonable variable ranges, and a structure similar to real data.

What is a possible useful implementation of secondary services for dental implant registry?

The results indicate that the stakeholders perceived most of the features we developed as valuable, allowing them to quickly explore various relationships in the data. The common theme among the most useful features was that they were directly related to the goals of a dental quality register and that they had a relatively low degree of complexity. The stakeholders who were best suited to use the features were professionals. The application will likely require additional adjustments to adapt it to a broader audience.

How can the process of generating synthetic data manually and creating secondary services contribute to the improvement of the data model used by a dental implant quality registry?

The parallel development of synthetic data and registry services may give stakeholders new insights regarding the data model's structure. The stakeholders involved in the project did not suggest new additions or modifications to the data model during the project, indicating that the parallel development did not directly contribute to improving the data model. However, it helped us, as developers, to discover suboptimal groupings in the data model.

Chapter 9

Future Work

Although we have explored the use of synthetic data to develop novel services, more work is needed to understand all aspects of this method. This understanding may be essential to learn more about the feasibility of the method and its use in other settings. We present some future work to investigate these aspects.

- It would be interesting to examine more closely how synthetic data and its realism affected the development process. Part of this investigation should involve the establishment of guidelines for the use of synthetic data and the establishment of measurements for the quality of synthetic data.
- How does the domain knowledge of the stakeholders and the type of application influence the use of synthetic data in the development process.
- We used mainly focus groups to evaluate the prototype. Feedback from these groups influenced the tasks for subsequent iterations. Future studies could compare evaluation methods to identify which are best suited to evaluate a prototype.

The prototype we developed is a potential starting point for dental implant registry services. Due to the lack of time, we have left many modifications and additions for the future. We have listed many of them in Table E.1. Future work should identify the best composition of features for a group of stakeholders. This collection may include features that we have already created and new features that we suggested or did not consider. New features may involve more overviews of the registry's data, patients, and clinicians' performance. An exciting feature might be to add statistical process control to monitor specific indicators over time, alerting if they exceed a threshold. Finding the appropriate use case and indicators would be essential for this process.

Another critical aspect before the prototype can be used in any real-life setting is security. We discuss this in Section 5.4. It is paramount that users cannot identify patients using the application. Therefore, developers must map any attack vectors adequately and consider other scenarios to ensure that this is impossible. Future developers should consider the aspects of authorisation and authentication more thoroughly. Access rights to view certain features may be denied to specific stakeholders. However, a potentially exciting approach would be more fine-grained, reducing access to particular options but allowing access to the feature. Future work should also consider how best to implement authentication to be part of the stakeholders' workflow.

We did not explore how to incorporate patient-reported data into the application. In healthcare, the importance of these data is increasing and some projects are trying to communicate them through dashboards (Lindblad et al., 2016). Therefore, future work should explore the possibilities of similar solutions within the domain of implant dentistry. This work could involve investigating what data should be recorded and how to visualise them to ensure informing and including patients in the care process.

Appendix A

Interview questions

1. Which dashboard features did you find useful? Why?
2. Which dashboard features were not useful? Why?
3. Which other features do you wish were added to the dashboard?
4. Were you bothered by the use of synthetic data in this evaluation? Why?
5. Was the synthetic data useful through the various iterations' evaluations?
6. How would the synthetic data's realism affect this usefulness.
7. See table B.1

Appendix B

Stakeholder Mapping

Table B.1: Which features are suitable for which stakeholders?

Dashboard pages	Stakeholders					
	Researchers	Registry adm.	Managers	Clinicians	Patients	Other
Removals						
Surv analysis						
Factor Surv.						
Clinic						
Explorer						
Analyses						
Report						
Std Report						

Appendix C

Answers

C.1 Statistician 1

Q1: Which dashboard features did you find useful? Why?

The overview of the clinic. This feature allows clinics to assess their performance and compare it with the national mean. Consequently, it may initiate an investigation into the clinic's performance to try and identify the reasons.

Q2: Which dashboard features were not useful? Why?

The Analyse functionality. Creating models for research purposes requires a lot of preparation. Models made with this feature cannot be used in research, although stakeholders can use the feature to identify areas of interest.

Q3: Which other features do you wish were added to the dashboard

The reporting functionality is an exciting feature, and developers can extend it using an automatic dispatch system.

Q4: Were you bothered by the use of synthetic data in this evaluation? Why?

It is impossible to know whether the data were real or false without domain knowledge.

Q5: Was the synthetic data useful through the various iterations' evaluations?

Absolutely. Very useful to demonstrate the features.

Q6: How would the synthetic data's realism affect this usefulness?

No, too little domain knowledge to know what real data would look like.

Table C.1: Q7: Which features are suitable for which stakeholders?

	Stakeholders					
Dash-board pages	Research-ers	Registry adm.	Managers	Clinicians	Patients	Other
Removals	X	X	X	X	X	X
Surv analysis	X	X	X	X	X	X
Factor Surv.	X	X	X	X	X	X
Clinic	X	X	X	X	X	X
Explorer	X	X	X	X	X	X
Analyses	X	X	X	X	X	X
Report	X	X	X	X	X	X
Std Report	X	X	X	X	X	X

C.2 Statistician 2

Q1: Which dashboard features did you find useful? Why?

- Overview removals, but not necessarily limited to removal reasons, but other overviews as well.
- Clinic page is informative and easy to understand.
- Factor survival. This feature may be more valuable than survival analysis due to the ease of comparing the survival rate of different factors. Moreover, the flexibility given by choosing comparison years is also valuable. This setting may be further enhanced by adding a third year.

Q2: Which dashboard features were not useful? Why?

More technical features that require special skills to interpret the results are less useful. For example, the "Analyze" feature. Some users may overinterpret the effect; therefore, access should be restricted.

Q3: Which other features do you wish were added to the dashboard

Statistical process control where some attributes are monitored over time, and if any values exceed a threshold, an alarm is activated.

Q4: Were you bothered by the use of synthetic data in this evaluation? Why?

Synthetic data help evaluate features because the focus remains on features, not data. If the data had been real, the statistician explained that he would be

Table C.2: Q7: Which features are suitable for which stakeholders?

Dashboard pages	Stakeholders					
	Researchers	Registry adm.	Managers	Clinicians	Patients	Other
Removals	X	X	X	X	X	X
Surv analysis	X	X	X	X	X	X
Factor Surv.	X	X	X	X	X	X
Clinic	X	X	X	X	X	X
Explorer	X	X	X	X	X	X
Analyses	X	X	X	X	X	X
Report	X	X	X	X	X	X
Std Report	X	X	X	X	X	X

more interested in interpreting what the charts showed and the implications of the results.

Q5: Was the synthetic data useful through the various iterations' evaluations?

Yes. Compared to real data, it is easy to avoid being distracted by trying to interpret the data. The synthetic data used in the iterations were reasonable.

Q6: How would the synthetic data's realism affect this usefulness?

Demonstrating some features, such as survival analysis, would have been less interesting if we had used random data. Now, the graph is recognisable and shows the feature.

C.3 Researcher 1

Q1: Which dashboard features did you find useful? Why?

Useful dashboard features can be used to identify characteristics with a higher probability of fault.

Q2: Which dashboard features were not useful? Why?

It is difficult to say which features were less valuable because each clinician and scientist could use this application to achieve different goals.

Table C.3: Q7: Which features are suitable for which stakeholders?

	Stakeholders					
Dash-board pages	Research-ers	Registry adm.	Managers	Clinicians	Patients	Other
Removals	X	X		X		
Surv analysis	X			X		
Factor Surv.	X			X		
Clinic	X			X		
Explorer	X	X		X		
Analyses	X					
Report			X	X		
Std Report			X			

Q3: Which other features do you wish were added to the dashboard

Not at the moment.

Q4: Were you bothered by the use of synthetic data in this evaluation? Why?

No, it is difficult to judge the realism of the data.

Q5: Was the synthetic data useful through the various iterations' evaluations?

It was exciting to watch the demonstrations, although the data shown in the presentations were synthetic.

Q6: How would the synthetic data's realism affect this usefulness?

It is useful to have some realism in the data. Completely random data would have been difficult to relate to.

C.4 Researcher 2

Q1: Which dashboard features did you find useful? Why?

It depends on whether you are a clinician or a scientist. Clinicians and researchers may benefit from information about removals. This information may be what implants are removed and why. However, survival analysis and other analyses may mostly benefit researchers, as you need some knowledge of statistics.

Q2: Which dashboard features were not useful? Why?

Analyses. This feature may be difficult to interpret and adapt to your needs. In addition, clinicians may find it easier to interpret graphs compared to analyses.

Q3: Which other features do you wish were added to the dashboard

A graph showing the cumulative growth of removals for a specific removal reason and more information about the chart and what is currently displayed in the application.

Q4: Were you bothered by the use of synthetic data in this evaluation? Why?

The scientist noticed some irregularities, such as many long implants, but was not distracted by them. It is challenging to get all the attributes of the synthetic data to be realistic.

Q5: Was the synthetic data useful through the various iterations' evaluations?

Yes, it would have been better with real data, but lacking such data the synthetic data proved helpful in demonstrating the application.

Q6: How would the synthetic data's realism affect this usefulness?

It would have been better to have actual data, but lacking such data, the synthetic data used in this project was an improvement compared to random data. We probably would not have accomplished as much if we had used entirely random data.

C.5 Clinician 1

Q1: Which dashboard features did you find useful? Why?

The most valuable aspect for me is how my patients are doing overall. Of course, this must be seen relative to other patients. If there are significant deviations in my performance, the application should have features that can help determine the cause. In other words, features that give an overview and allow a closer inspection of the data.

Q2: Which dashboard features were not useful? Why?

It was difficult to select features that were not useful because much depends on the area of use.

Q3: Which other features do you wish were added to the dashboard

Any feature that allows clinicians to compare their performance with others or their previous performance. For example, the survival rate of tooth implants inserted by a particular clinician compared to other clinicians using the same implants.

Table C.4: Q7: Which features are suitable for which stakeholders?

	Stakeholders					
Dashboard pages	Researchers	Registry adm.	Managers	Clinicians	Patients	Other
Removals	X	X	X	X		
Surv analysis	X	X				
Factor Surv.	X	X	X	X		
Clinic	X	X	X	X		
Explorer		X		X		
Analyses	X	X	X			
Report	X			X		
Std Report		X				

Q4: Were you bothered by the use of synthetic data in this evaluation? Why?

No, I did not study the data, only the features were in focus during this session.

Q5: Was the synthetic data useful through the various iterations' evaluations?

Only

Q6: How would the synthetic data's realism affect this usefulness?

The clinician said that the realism of the synthetic data did not affect him in the evaluation. However, he adds that it might have been distracting if we had not informed him beforehand.

Table C.5: Q7: Which features are suitable for which stakeholders?

Dashboard pages	Stakeholders: Researchers	Registry adm.	Managers	Clinicians	Patients	Other
Removals	X	X	X	X	X	
Surv analysis	X	X	X	X	X	
Factor Surv.	X	X		X		
Clinic		X		X	X	
Explorer			X	X		
Analyses	X	X				
Report		X	X	X		
Std Report		X	X	X		

Appendix D

Think aloud

We summarise the recordings of the think-aloud session with stakeholders focused on feedback related to the application and the features. Because the sessions were held in Norwegian, questions and answers were translated. The summaries are grouped according to stakeholders and features.

D.1 Statistician 1

Removals

- It would have been helpful to know the range of different variables before/after selecting them in the options.
- It is convenient that the application shows the number of observations per removal reason in the bar chart.
- The application changes the colours of the bar graph for removal reasons when filtering. This recolouring is not optimal.
- The advanced option produces charts that may overwhelm users, but may be informative in some cases. However, filtering the faceting rows reduces the complexity.

Surv analysis

- The possibility of zooming in on the graph would have been helpful.
- The survival analysis is potentially valuable for quickly reviewing whether a particular factor influences the survival of a tooth implant.

Factor survival

- The statistician notes that he does not have enough domain knowledge to evaluate the feature's usefulness; however, the feature appears clear and informative.

Clinic

- This feature is the same as what our department will create. We will use indicators to highlight a clinic's performance and show changes over time.
- It would have been nice to click on the quality indicators to get more information on the clinic's performance. For example, if the percentage of insertion complications was high, clicking on the box could show the complication reasons.
- Remove decimals on the chart showing which implants were removed or inserted.

Explorer

- It is easy to filter the data set by selecting removals or insertions.
- The advanced Explorer plot was interesting, allowing users to create almost any graph.

Analyses

- It is essential to inform users that any results of this feature are indicative only.

D.2 Statistician 2

Removals

- The application seems dynamic with many settings. Other registries may have more static visualisation, but this application allows one to change many aspects of the charts.
- It would take some time to understand how the options are structured.
- No possibility to view all the years combined in a single graph.
- This feature provides a valuable overview that stakeholders could have used directly in an annual report.
- It would have been convenient to have some functionality with the option to save graphs and view previously saved graphs. This feature would make it easier to retrieve graphs used in annual reports.
- After creating a large graph showing the reasons for removing each tooth position, the statistician suggests adding a more crude grouping. For example, group removals by quadrant or upper/lower jaw.
- All the feedback is just nitpicking. Most of the functionality is overall very good.

Surv analysis

- Remove crosses from the graph that mark when an implant fails. This removal will ensure a much clearer and cleaner graph.

- It is interesting to create such figures to explore the data quickly.
- It is easy to change the study duration with the slider. This attribute is usually cumbersome to modify in other applications such as STATA. A similar option should have been available for the y-axis.
- All nuances are not taken into account; for example, if a person has recently died, we would censor their implants because we would not gain new information on their implants. Connecting the application directly to the National Population Register would have ensured updated data on this, but such functionality is long in the future.

Factor survival

- Informative feature.
- After comparing the survival rates of implants grouped by clinics, the statistician notes that there may be thousands of clinics. Therefore, some consideration should be made on how to view only a subgroup of these and perhaps group them by county and region.

Clinic

- Excellent overview with funnel plots and time series. It would have been helpful to have a report generator that could easily create such an overview in a report.

Explorer

- Very helpful with the easy mode and the advanced mode. It is easy to become overwhelmed by too many options. The Easy mode shows you what you usually want to investigate, while the Advanced mode allows you to go in-depth.
- Maybe include the possibility to choose which graph type to use.

Analyses

- GLM gives you only a linear prediction. Hence, GAM may be more useful.

Report

- A automatic send out each every six months would have been nice.
- The statistician mentions a feature in which the dashboard landing page lists various graphs, allowing the user to select which charts to use in a report. This feature is more convenient than going through each page and clicking "Add to report". Additionally, such a feature would be even more useful if it remembered the last time setting, as it is challenging to remember what you did six months ago.
- When creating a report, it is essential to know which global filters were applied to the graph. This feature is currently not implemented.

D.3 Researcher 1

Removals

- The researcher said that it would have been interesting to view the reasons for implant removal in groups such as front teeth. This grouping was partly possible but was hidden under advanced options. Therefore, the researcher needed some guidance to locate the options.

Surv analysis

- The researcher asks some questions about how to interpret the chart and whether there is additional information.
- After selecting the implant name as an independent variable, the researcher revisits the topic of grouping variables. The researcher notes that grouping the implant names according to characteristics would have been interesting. There is some possibility to do this grouping, but the options are not easy to discover.

Factor survival

- After selecting the tooth positions as a factor, the researcher asks questions about how to filter the positions.
- The researcher comments that the functionality offered by this feature seems valuable.
- The researcher needs some information to understand the drill-down options.

Clinic

- The researcher notes that clinics should only have access to their clinic and asks if it is possible to compare the clinic's performance with the mean.
- In the application, all clinics are placed in the funnel plot and, although only your clinic is named, it may be possible to recognise other clinics. The researcher states that it is vital that clinics do not feel stigmatised, which should be considered when using such visualisations.

Explorer

- The researcher notes that, in simple mode, the factors along the x-axis are repeated in the legend.
- The researcher explores the distribution of insertions grouped by tooth positions, commenting that the data seem reasonable because molars are frequently inserted.
- The simple mode gives an interesting overview of the distribution.
- Although the layout of the visualisations and the options is logical, a user manual would have been helpful.

D.4 Researcher 2

Removals

- Text that more clearly describes the features currently shown in the application. For example, the removal reason ”external trauma” can be a removal reason and a reason for missing a tooth. Therefore, the application must clearly distinguish what is displayed on the screen.
- When using the advanced options, the researcher created a graph that contained only NA. Even though such graphs can occur due to the flexibility given to users, they might be confusing.
- A graph showing how cumulative growth increases over time for a specific removal reason would have been interesting.
- The researcher found it challenging to know which options were grouped. An example of such a group is the selection of facet rows in the advanced options. After selecting the faceting row, the application will display another input box below, allowing users to filter the rows.

Surv analysis

- A clever feature.
- No information on the scale of the x-axis.
- After performing a survival analysis with the independent variable set to Antibiotics, the researcher notes that here a global filter could be beneficial to compare similar patients.

Factor survival

- A grouping factor used in the application is the length of the implant. This factor can be used to compare the survival rate. However, one of these groups is defined as an implant length greater than 15 mm. According to the researcher, this grouping seems strange because it includes very long implants.
- It might be challenging to use the drill-down function due to the enormous amount of implants. The attributes of the implants are more relevant to explore.

Clinic

- The researchers have some questions about how to interpret the different features.
- After testing the global filter and selecting bone augmentation, the researcher remarks that it might be interesting to differentiate between bone from the hip and artificial bone.

Explorer

- More information on what is displayed on the graph.

Analyses

- This feature is for users who want to do research.

D.5 Clinician 1

Options

When demonstrating the options and the global filter, we side-tracked the evaluation with a discussion about the problem of comparing groups directly.

- The clinician explained that if the goal is to identify the performance of tooth implants, there are many parameters related to the patients that will affect their success rate. Therefore, the tooth implant may not be the problem but the patient's condition. The same is true when comparing the performance of clinics. Some clinics specialise in complex cases, resulting in a lower overall success rate. Another example is that clinicians may use specific implants under challenging circumstances, resulting in a low survival rate. Hence, these implants may appear to underperform but may actually be good implants under these conditions. Examples of such relationships are critical to know when interpreting data from the tooth implant registry
- Another factor to be aware of is that clinicians may have a varying degree of understanding or knowledge of the terminology used by the registry, causing uncertainty in the data they register. For example, if the registry uses common names for certain antibiotics, some clinicians may not know how to report a specific antibiotic. Another example is complications; one clinician might define an event as a complication and another might ignore it.
- Of course, if the numbers are large enough, maybe all these concerns are unfounded.

Removals

- As an example, we selected the factor "vendor" in the advanced options. The graph produced showed how the distribution of removals reasons differed by years and vendors. The clinician first responded that such visualisations were interesting because determining the underperforming implants is one of the main objectives of the registry. However, he added that grouping by "vendor" is not that interesting because of the enormous amount of implants related to each. Consequently, each vendor can have high-quality and low-quality tooth implants.
- As another example, we selected "position" in the advanced options, producing a graph displaying the distribution of removal reasons by years and insertion positions. The clinician agreed that investigating factors unrelated to specific implants, such as position, was also interesting.
- After showing the clinician all the possible options for creating the graphs, he commented on the terms the application used. We had taken these directly from the data model. However, the clinician remarked that they

were in English and not always consistent with the terminology used in daily practise.

- The principle of defining graphs with advanced options is helpful. The registry needs to collect a significant amount of data before this feature is useful.

Surv analysis

- The feature may be helpful, but the analysis is dependent on the quality of the data.
- It is interesting to examine LotNr with this feature.
- This tool is best suited for researchers, but it can be interesting for clinicians.
- The area of use for clinicians will not be to determine which implants to use for a particular patient. However, clinicians may use it to obtain information on implants in general. Of course, this usage requires years of data collection.

Factor survival

- This feature is high quality from a technical point of view.
- This feature will probably be the most widely used by researchers.
- Only a limited number of clinicians would probably use this feature.
- It is fascinating to explore the data using this feature, as it provides excellent information.
- After again viewing all the factors collected by the registry in the options, the clinician notes that he fears that registering all this information would be too cumbersome for most dentists as part of their daily routine when inserting an implant. Unfortunately, there may be a need for some force to ensure that all clinicians register these data.

Clinic

- The clinician asks whether you can choose which quality indicators the application displays. This option is currently not implemented. Therefore, the clinician suggests that having the ability to view other summary statistics would have been interesting and an overall survival rate of tooth implants
- It is interesting to compare your clinic with others, as demonstrated in the funnel graph. Clinicians who work in a clinic must only be able to access their clinic and not other clinics.
- Comparing a clinic with other clinics can act as a motivation to improve.
- The application currently shows the overall survival rate after two years in the funnel plot. The clinician suggests allowing users to modify the funnel plot, exploring the survival rate related to other factors.

Explorer

- The overview of the distribution of different variables is a practical feature, but the main focus should be on the implants that clinicians remove.

Appendix E

Source code

The source code for the dashboard is available at this URL: `https://github.com/oddhus/DashboardDirect`. The source code for the synthetic data is available at this URL: `https://github.com/oddhus/GeneratePatientData`

Table E.1: List of changes and additions suggested during interviews and think-alouds

ID	Synthesis
CHANGES	
R2	More information about the chart being displayed.
S1	Consistent chart colouring.
S1	Ability to zoom in on graphs.
S1	In "Clinic Information" get more information about quality indicators by clicking on the infoboxes.
R2	Improve the grouping of options. It is difficult to understand what filters belong to particular options.
S2	Ability to change the grouping of variables form crude to fine, e.g. from upper/lower jaw to tooth position.
S2	Remove crosses from graph in "Survival Analysis".
C1	Possibility to explore other factors in the funnel plot on the "Clinic information" page.
ADDITIONS	
S1	Add an automatic dispatch system to the reporting functionality.
S2	Statistical process control, monitoring attributes over time.
R2	A graph showing the cumulative growth of removals for a specific removal reason.
C1	Any feature that allows clinicians to compare their performance with others or their previous performance.
S1	Ability to inspect variable ranges before/after selecting them in options.
S2	Improve reporting functionality, cumbersome to click add and navigate.
R1	Add functionality which may assist the user if stuck.

Bibliography

Adami, H.-O. and M A Hernán (2015). "Learning how to improve healthcare delivery: the Swedish Quality Registers." In: *Journal of Internal Medicine* 277.1, pp. 87–89. DOI: https://doi.org/10.1111/joim.12315. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/joim.12315.

Ahmad, Nabeel and Najeeb Saad (2012). "Effects of Antibiotics on Dental Implants: A Review." In: *Journal of Clinical Medicine Research* 4.1, p. 1. DOI: 10.4021/JOCMR658W. URL: /pmc/articles/PMC3279494/%20/pmc/articles/PMC3279494/?report=abstract%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3279494/.

Althar, Raghavendra Rao and Debabrata Samanta (Mar. 2021). "The realist approach for evaluation of computational intelligence in software engineering." In: *Innovations in Systems and Software Engineering* 17.1, pp. 17–27. ISSN: 16145054. DOI: 10.1007/S11334-020-00383-2/FIGURES/2. URL: https://link.springer.com/article/10.1007/s11334-020-00383-2.

Alvertis, Iosif et al. (Jan. 2016). "User Involvement in Software Development Processes." In: *Procedia Computer Science* 97, pp. 73–83. ISSN: 1877-0509. DOI: 10.1016/J.PROCS.2016.08.282.

Bakken, Inger Johanne et al. (2020). "The Norwegian Patient Registry and the Norwegian Registry for Primary Health Care: Research potential of two nationwide health-care registries." In: *Scandinavian Journal of Public Health* 48.1, pp. 49–55. DOI: 10.1177/1403494819859737. URL: https://doi.org/10.1177/1403494819859737.

Bangor, Aaron et al. (Aug. 2008). "An Empirical Evaluation of the System Usability Scale." In: *https://doi.org/10.1080/10447310802205776* 24.6, pp. 574–594. ISSN: 10447318. DOI: 10.1080/10447310802205776. URL: https://www.tandfonline.com/doi/abs/10.1080/10447310802205776.

Batalden, Maren et al. (July 2016). "Coproduction of healthcare service." In: *BMJ Quality & Safety* 25.7, p. 509. DOI: 10.1136/BMJQS-2015-004315. URL: /pmc/articles/PMC4941163/%20/pmc/articles/PMC4941163/?report=abstract%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4941163/.

Benoit, Ballester et al. (Apr. 2022). "Current state of dental informatics in the field of health information systems: a scoping review." In: *BMC Oral Health 2022 22:1* 22.1, pp. 1–17. ISSN: 1472-6831. DOI: 10.1186/S12903-022-02163-9. URL: https://bmcoralhealth.biomedcentral.com/articles/10.1186/s12903-022-02163-9.

Beschnidt, Sven Marcus et al. (Dec. 2018). "Implant success and survival rates in daily dental practice: 5-year results of a non-interventional study using CAMLOG SCREW-LINE implants with or without platform-switching abutments." In: *International Journal of Implant Dentistry* 4.1. DOI: 10.1186/S40729-018-0145-3. URL: /pmc/articles/PMC6212375/%20/pmc/articles/PMC6212375/?report=abstract%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6212375/.

Bourque, Pierre et al. (2014). *Guide to the Software Engineering Body of Knowledge (SWEBOK(R)): Version 3.0.* 3rd. Washington, DC, USA: IEEE Computer Society Press. ISBN: 0769551661.

Buser, Daniel et al. (2017). "Modern implant dentistry based on osseointegration: 50 years of progress, current trends and open questions." In: *Periodontology 2000* 73.1, pp. 7–21. DOI: https://doi.org/10.1111/prd.12185. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/prd.12185.

Charles, Dustin et al. (2015). "Adoption of Electronic Health Record Systems among U.S. Non-Federal Acute Care Hospitals: 2008-2014." In: 23.

Chaudhry, Basit et al. (May 2006). "Systematic review: Impact of health information technology on quality, efficiency, and costs of medical care." In: *Annals of Internal Medicine* 144.10, pp. 742–752. ISSN: 00034819. DOI: 10.7326/0003-4819-144-10-200605160-00125/SUPPL{_}FILE/CHAUDHRY{_}AT2{_}144-10-742-DC1-1.PDF.

Chen, Richard J. et al. (June 2021). "Synthetic data in machine learning for medicine and healthcare." In: *Nature Biomedical Engineering 2021 5:6* 5.6, pp. 493–497. ISSN: 2157-846X. DOI: 10.1038/s41551-021-00751-8. URL: https://www.nature.com/articles/s41551-021-00751-8.

Chuang, S. K. et al. (Aug. 2002). "Risk factors for dental implant failure: A strategy for the analysis of clustered failure-time observations." In: *Journal of Dental Research* 81.8, pp. 572–577. ISSN: 00220345. DOI: 10.1177/154405910208100814. URL: https://journals.sagepub.com/doi/10.1177/154405910208100814?url_ver=Z39.88-2003&rfr_id=ori%3Arid%3Acrossref.org&rfr_dat=cr_pub++0pubmed.

Clarke, S et al. (2016). "Using Dashboard Technology and Clinical Decision Support Systems to Improve Heart Team Efficiency and Accuracy: Review of the Literature." In: *Stud Health Technol Inform* 225, pp. 364–366. ISSN: 0926-9630.

Cochran, David et al. (June 2007). "Clinical Field Trial Examining an Implant With a Sand-Blasted, Acid-Etched Surface." In: *Journal of Periodontology* 78.6, pp. 974–982. ISSN: 1943-3670. DOI: 10.1902/JOP.2007.060294. URL: https://onlinelibrary.wiley.com/doi/full/10.1902/jop.2007.060294%20https://onlinelibrary.wiley.com/doi/abs/10.1902/jop.2007.060294%20https://aap.onlinelibrary.wiley.com/doi/10.1902/jop.2007.060294.

Concannon, David et al. (Apr. 2019). "Developing a Data Dashboard Framework for Population Health Surveillance: Widening Access to Clinical Trial Findings." In: *JMIR formative research* 3.2. ISSN: 2561-326X. DOI: 10.2196/11342. URL: https://pubmed.ncbi.nlm.nih.gov/30946016/.

Dagliati, Arianna et al. (Aug. 2018). "A dashboard-based system for supporting diabetes care." In: *Journal of the American Medical Informatics Association* 25.5, pp. 538–547. ISSN: 1527-974X. DOI: 10.1093/jamia/ocx159. URL: https://doi.org/10.1093/jamia/ocx159.

Davis, Alan M. (1992). "Operational Prototyping: A New Development Approach." In: *IEEE Software* 9.5, pp. 70–78. ISSN: 07407459. DOI: 10.1109/52.156899.

De Croon, Robin et al. (Dec. 2015). “Design and evaluation of an interactive proof-of-concept dashboard for general practitioners.” In: *Proceedings - 2015 IEEE International Conference on Healthcare Informatics, ICHI 2015*, pp. 150–159. DOI: 10.1109/ICHI.2015.25.

Dilmegani, Cem. “The Ultimate Guide to Synthetic Data: Uses, Benefits & Tools.” URL: https://research.aimultiple.com/synthetic-data/. *Last retrieved: 13.01.2022*.

Dolan, James G. et al. (2013). “Development and initial evaluation of a treatment decision dashboard.” In: *BMC Medical Informatics and Decision Making* 13.1, p. 51. ISSN: 14726947. DOI: 10.1186/1472-6947-13-51. URL: /pmc/articles/PMC3639808/%20/pmc/articles/PMC3639808/?report=abstract%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3639808/.

Dowding, Dawn et al. (2015). “Dashboards for improving patient care: Review of the literature.” In: *International Journal of Medical Informatics* 84.2, pp. 87–100. ISSN: 1386-5056. DOI: https://doi.org/10.1016/j.ijmedinf.2014.10.001. URL: https://www.sciencedirect.com/science/article/pii/S1386505614001890.

Dresch, Aline et al. (2015). “Design science research.” In: *Design science research*. Springer, pp. 67–102.

Dunn, Sandra et al. (2016). “A mixed methods evaluation of the maternal-newborn dashboard in Ontario: dashboard attributes, contextual factors, and facilitators and barriers to use: a study protocol.” In: *Implementation science : IS* 11, p. 59. ISSN: 1748-5908. DOI: 10.1186/s13012-016-0427-1. URL: https://pubmed.ncbi.nlm.nih.gov/27142655%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4855363/.

Elias, Carlos Nelson (2011). “Factors affecting the success of dental implants.” In: *Implant dentistry: a rapidly evolving practice. Rijeka: InTech*, pp. 319–364. URL: https://books.google.com/books?hl=no&lr=&id=S9CPDwAAQBAJ&oi=fnd&pg=PA319&dq=success+dental+implants&ots=z1AtTVbnQb&sig=Eti2u757svW1jDlFASyOFXMgXxg.

Escalona, María et al. (Aug. 2021). “Don’t Throw your Software Prototypes Away. Reuse them!” In: *International Conference on Information Systems Development (ISD)*. URL: https://aisel.aisnet.org/isd2014/proceedings2021/hci/2.

Evergreen, Stephanie and Chris Metzner (Dec. 2013). “Design Principles for Data Visualization in Evaluation.” In: *New Directions for Evaluation* 2013.140, pp. 5–20. ISSN: 10976736. DOI: 10.1002/EV.20071.

Few, Stephen (2006). *Information dashboard design: The effective visual communication of data*. O’Reilly Media, Inc. ISBN: 0596100167.

Finkelstein, Joseph et al. (Mar. 2020). “Using big data to promote precision oral health in the context of a learning healthcare system.” In: *Journal of Public Health Dentistry* 80.S1, S43–S58. ISSN: 1752-7325. DOI: 10.1111/JPHD.12354. URL: https://onlinelibrary.wiley.com/doi/full/10.1111/jphd.12354%20https://onlinelibrary.wiley.com/doi/abs/10.1111/jphd.12354%20https://onlinelibrary.wiley.com/doi/10.1111/jphd.12354.

Fiskeseth, Elise (2022). “Developing a User Interface and Integrated Workflow Data Collection for a Dental Implant Quality Register (In submission).” In: *Master thesis in Software Engineering, Department of Informatics, University of Bergen, Norway.*

Franklin, Amy et al. (2017). “Dashboard visualizations: Supporting real-time throughput decision-making.” In: *Journal of Biomedical Informatics* 71, pp. 211–221. ISSN: 1532-0464. DOI: https://doi.org/10.1016/j.jbi.2017.05.024. URL: https://www.sciencedirect.com/science/article/pii/S1532046417301235.

Fredriksson, Mio et al. (Aug. 2017). “Are data from national quality registries used in quality improvement at Swedish hospital clinics?” In: *International Journal for Quality in Health Care* 29.7, pp. 909–915. ISSN: 1353-4505. DOI: 10.1093/intqhc/mzx132. URL: https://doi.org/10.1093/intqhc/mzx132.

Furnes, Ove et al. (2020). “Nasjonalt Register for Leddproteser Årsrapport for 2020 med plan for forbedringstiltak Fagrådet.” In:

Furu, Kari et al. (Feb. 2010). “The Nordic Countries as a Cohort for Pharmacoepidemiological Research.” In: *Basic & Clinical Pharmacology & Toxicology* 106.2, pp. 86–94. ISSN: 1742-7843. DOI: 10.1111/J.1742-7843.2009.00494.X. URL: https://onlinelibrary.wiley.com/doi/full/10.1111/j.1742-7843.2009.00494.x%20https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1742-7843.2009.00494.x%20https://onlinelibrary.wiley.com/doi/10.1111/j.1742-7843.2009.00494.x.

Ghazisaeidi, Marjan et al. (2015). “Development of Performance Dashboards in Healthcare Sector: Key Practical Issues.” In: *Acta Informatica Medica* 23.5, p. 317. DOI: 10.5455/AIM.2015.23.317-321. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4639357/.

Goldfeld, Keith and Jacob Wujciak-Jens (2020). “simstudy.” In: *Last retrieved 16.09.21* 5.54, p. 2763. URL: https://doi.org/10.21105/joss.02763.

Goodfellow, Ian et al. (June 2014). “Generative Adversarial Networks.” In: *Communications of the ACM* 63.11, pp. 139–144. ISSN: 15577317. DOI: 10.1145/3422622. URL: https://arxiv.org/abs/1406.2661v1.

Gordon, V. Scott and James M. Bieman (1995). “Rapid Prototyping: Lessons Learned.” In: *IEEE Software* 12.1, pp. 85–95. ISSN: 07407459. DOI: 10.1109/52.363162.

Hahner, Martin et al. (2019). “Semantic Understanding of Foggy Scenes with Purely Synthetic Data.” In:

Hartmann-Johnsen, Olaf Johan et al. (2019). “Using clinical cancer registry data for estimation of quality indicators: Results from the Norwegian breast cancer registry.” In: *International Journal of Medical Informatics* 125, pp. 102–109. ISSN: 1386-5056. DOI: https://doi.org/10.1016/j.ijmedinf.2019.03.004. URL: https://www.sciencedirect.com/science/article/pii/S1386505618307664.

Hartzler, Andrea L. et al. (Mar. 2015). “Integrating Patient-Reported Outcomes into Spine Surgical Care through Visual Dashboards: Lessons Learned from Human-Centered Design.” In: *eGEMs* 3.2, p. 2. ISSN: 2327-9214. DOI: 10.13063/2327-9214.1133. URL: /pmc/articles/PMC4431498/%20/pmc/articles/PMC4431498/?report=abstract%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4431498/.

Hevner, Alan R (2007). “A three cycle view of design science research.” In: *Scandinavian journal of information systems* 19.2, p. 4. URL: https://aisel.aisnet.org/sjis/vol19/iss2/4.

Hevner, Alan R and Samir Chatterjee (2010). “Design Science Research in Information Systems.” In: pp. 9–22. DOI: 10.1007/978-1-4419-5653-8{_}2.

URL: https://link.springer.com/chapter/10.1007/978-1-4419-5653-8_2.
Hove, Siw Elisabeth and Bente Anda (2005). "Experiences from Conducting Semi-Structured Interviews in Empirical Software Engineering Research." In:
"Interact. Analyze. Communicate" (n.d.). In: *Last retrieved: 23.08.21* (). URL: https://shiny.rstudio.com/.
James, Stefanie et al. (Dec. 2021). "Synthetic data use: exploring use cases to optimise data utility." In: *Discover Artificial Intelligence 2021 1:1* 1.1, pp. 1–13. ISSN: 2731-0809. DOI: 10.1007/S44163-021-00016-Y. URL: https://link.springer.com/article/10.1007/s44163-021-00016-y.
Jenny, Gregor et al. (2016). "A systematic review and meta-analysis on the influence of biological implant surface coatings on periimplant bone formation." In: *Journal of Biomedical Materials Research Part A* 104.11, pp. 2898–2910. DOI: https://doi.org/10.1002/jbm.a.35805. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/jbm.a.35805.
Jernberg, Tomas et al. (2014). "SWEDEHEART Annual Report 2013." In: *Matador Kommunikation AB Uppsala.*
Joda, Tim et al. (Nov. 2018). "Population-Based Linkage of Big Data in Dental Research." In: *International Journal of Environmental Research and Public Health* 15.11. DOI: 10.3390/IJERPH15112357. URL: /pmc/articles/PMC6265733/%20/pmc/articles/PMC6265733/?report=abstract%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6265733/.
John Brooke (June 1996). "SUS: A 'Quick and Dirty' Usability Scale." In: *Usability Evaluation In Industry*, pp. 207–212. DOI: 10.1201/9781498710411-35. URL: https://www.taylorfrancis.com/chapters/edit/10.1201/9781498710411-35/sus-quick-dirty-usability-scale-john-brooke.
Jyotsna. "The Power of R – And Why it's an Essential Skill for Data Analysts." URL: https://www.jigsawacademy.com/the-power-of-r-and-why-its-an-essential-skill-for-data-analysts/. *Last retrieved: 25.11.21.*
Khairat, Saif Sherif et al. (May 2018). "The Impact of Visualization Dashboards on Quality of Care and Clinician Satisfaction: Integrative Literature Review." In: *JMIR Hum Factors 2018;5(2):e22 https://humanfactors.jmir.org/2018/2/e22* 5.2, e9328. ISSN: 22929495. DOI: 10.2196/HUMANFACTORS.9328. URL: https://humanfactors.jmir.org/2018/2/e22.
Kieler, Helle (2010). "The Nordic health registers – an important source when evaluating the safety of antidepressants during pregnancy." In: *Clinical Epidemiology* 2.1, p. 205. ISSN: 11791349. DOI: 10.2147/CLEP.S10426. URL: /pmc/articles/PMC2964074/%20/pmc/articles/PMC2964074/?report=abstract%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2964074/.
Kinnaird, Peter and Mario Romero (2010). "Focus Groups for Functional InfoVis Prototype Evaluation: A Case Study." In:
Klinge, Björn, Mats Lundström, et al. (2018). "Dental Implant Quality Register—A possible tool to further improve implant treatment and outcome." In: *Clinical Oral Implants Research* 29.S18, pp. 145–151. DOI: https://doi.org/10.1111/clr.13268. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/clr.13268.
Klinge, Björn, Mariano Sanz, et al. (2018). "Dental implant register: Summary and consensus statements of group 2. The 5th EAO Consensus Conference 2018." In: *Clinical Oral Implants Research* 29.S18, pp. 157–159. DOI: https:

//doi.org/10.1111/clr.13269. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/clr.13269.

Kordon, Fabrice and Luqi (Sept. 2002). "An introduction to rapid system prototyping." In: *IEEE Transactions on Software Engineering* 28.9, pp. 817–821. ISSN: 00985589. DOI: 10.1109/TSE.2002.1033222.

Kruse, Clemens Scott and Amanda Beane (Feb. 2018). "Health Information Technology Continues to Show Positive Effect on Medical Outcomes: Systematic Review." In: *J Med Internet Res 2018;20(2):e41 https://www.jmir.org/2018/2/e41* 20.2, e8793. ISSN: 14388871. DOI: 10.2196/JMIR.8793. URL: https://www.jmir.org/2018/2/e41.

Kusiak, John W. and Martha Somerman (Aug. 2016). "Data science at the National Institute of Dental and Craniofacial Research: Changing dental practice." In: *The Journal of the American Dental Association* 147.8, pp. 597–599. ISSN: 0002-8177. DOI: 10.1016/J.ADAJ.2016.06.005.

"Kvalitetsforbedring." URL: https://www.kvalitetsregistre.no/kvalitetsforbedring. *Last retrieved: 31.01.22*.

Lee, Jae-Hoon et al. (2005). "Effect of implant size and shape on implant success rates: A literature review." In: *The Journal of Prosthetic Dentistry* 94.4, pp. 377–381. ISSN: 0022-3913. DOI: https://doi.org/10.1016/j.prosdent.2005.04.018. URL: https://www.sciencedirect.com/science/article/pii/S002239130500243X.

Lemos, Cleidiel Aparecido Araujo et al. (Apr. 2016). "Short dental implants versus standard dental implants placed in the posterior jaws: A systematic review and meta-analysis." In: *Journal of Dentistry* 47, pp. 8–17. ISSN: 0300-5712. DOI: 10.1016/J.JDENT.2016.01.005.

Leung, Joshua and Andy Cockburn (Dec. 2020). "An Empirical Evaluation of Collapsible Panel Interfaces." In: *2020 IEEE Asia-Pacific Conference on Computer Science and Data Engineering, CSDE 2020*. DOI: 10.1109/CSDE50874.2020.9411552.

Lewis, Clayton and John Rieman. "Task-Centered user interface design: A practical introduction." URL: https://hcibib.org/tcuid/tcuid.pdf. *Retrieved: 09.03.2022*.

Lindblad, S et al. (Aug. 2016). "Creating a culture of health: evolving healthcare systems and patient engagement." In: *QJM: An International Journal of Medicine* 110.3, pp. 125–129. ISSN: 1460-2725. DOI: 10.1093/qjmed/hcw188. URL: https://doi.org/10.1093/qjmed/hcw188.

Ludvigsson, Jonas F et al. (2019). "Swedish Inflammatory Bowel Disease Register (SWIBREG) – a nationwide quality register." In: *Scandinavian Journal of Gastroenterology* 54.9, pp. 1089–1101. DOI: 10.1080/00365521.2019.1660799. URL: https://doi.org/10.1080/00365521.2019.1660799.

Lusk, Edward J. and Michael Kersnick (Aug. 1979). "The Effect of Cognitive Style and Report Format on Task Performance: The MIS Design Consequences." In: *http://dx.doi.org/10.1287/mnsc.25.8.787* 25.8, pp. 787–798. ISSN: 00251909. DOI: 10.1287/MNSC.25.8.787. URL: https://pubsonline.informs.org/doi/abs/10.1287/mnsc.25.8.787.

Mahajan, Shiwani et al. (Jan. 2021). "Trends and Predictors of Use of Digital Health Technology in the United States." In: *The American Journal of Medicine* 134.1, pp. 129–134. ISSN: 0002-9343. DOI: 10.1016/J.AMJMED.2020.06.033.

Mannino, Miro and Azza Abouzied (Oct. 2019). "Is this real? Generating synthetic data that looks real." In: *UIST 2019 - Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology*, pp. 549–561. DOI: 10.1145/3332165.3347866. URL: https://doi.org/10.1145/3332165.3347866.

Maqbool, Bilal and Sebastian Herold (2021). "Challenges in developing software for the swedish healthcare sector." In: *HEALTHINF 2021 - 14th International Conference on Health Informatics; Part of the 14th International Joint Conference on Biomedical Engineering Systems and Technologies, BIOSTEC 2021*, pp. 175–187. DOI: 10.5220/0010248901750187.

Marieke McCloskey. "Task Scenarios for Usability Testing." URL: https://www.nngroup.com/articles/task-scenarios-usability-testing/. *Last retrieved: 10.01.2022*.

Marinho, Marcelo et al. (Dec. 2014). "A Systematic Review of Uncertainties in Software Project Management." In: *International Journal of Software Engineering & Applications* 5.6, pp. 1–21. DOI: 10.5121/ijsea.2014.5601. URL: http://arxiv.org/abs/1412.3690%20http://dx.doi.org/10.5121/ijsea.2014.5601.

Marini, Federico and Harald Binder (Oct. 2017). "Development of Applications for Interactive and Reproducible Research: a Case Study." In: *Genomics and Computational Biology* 3.1, e39–e39. ISSN: 2365-7154. DOI: 10.18547/GCB.2017.VOL3.ISS1.E39. URL: https://www.genomicscomputbiol.org/ojs3/GCB/article/view/21.

Merrill, Mallory. "What is Rapid Prototyping And Why is it Used in Development?" URL: https://devsquad.com/blog/what-is-rapid-prototyping-and-why-is-it-used-in-development/. *Retrieved: 02.05.22*.

Meskó, Bertalan et al. (Sept. 2017). "Digital health is a cultural transformation of traditional healthcare." In: *mHealth* 3, pp. 38–38. DOI: 10.21037/MHEALTH.2017.08.07. URL: /pmc/articles/PMC5682364/%20/pmc/articles/PMC5682364/?report=abstract%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5682364/.

Morgan, D. L. (1997). *Focus Groups as Qualitative Research: PLANNING AND RESEARCH DESIGN FOR FOCUS GROUPS*, pp. 32–46. ISBN: 9780761903437. URL: http://srmo.sagepub.com/view/focus-groups-as-qualitative-research/SAGE.xml.

Naemi, Roya et al. (2021). "Dental implant quality registries and databases: A systematic review." In: *Journal of Education and Health Promotion* 10.1, p. 214. ISSN: 2277-9531. DOI: 10.4103/jehp.jehp{_}1302{_}20. URL: https://www.jehp.net/article.asp?issn=2277-9531;year=2021;volume=10;issue=1;spage=214;epage=214;aulast=Naemi.

Nelson, Scott D. et al. (Dec. 2016). "A case report of refining user requirements for a health information exchange dashboard." In: *Applied Clinical Informatics* 7.1, pp. 22–32. ISSN: 18690327. DOI: 10.4338/ACI-2015-07-CR-0091/ID/BR0091-44. URL: http://www.thieme-connect.com/products/ejournals/html/10.4338/ACI-2015-07-CR-0091%20http://www.thieme-connect.de/DOI/DOI?10.4338/ACI-2015-07-CR-0091.

Neumann, Ana et al. (Sept. 2017). "Evaluating quality of dental care among patients with diabetes: Adaptation and testing of a dental quality measure in electronic health records." In: *The Journal of the American Dental Associa-*

tion 148.9, pp. 634–643. ISSN: 0002-8177. DOI: 10.1016/J.ADAJ.2017.04.017.

Oliver, Brant J et al. (2019). "Turning Feed-forward and Feedback Processes on Patient-reported Data into Intelligent Action and Informed Decision-making: Case Studies and Principles." In: *Medical Care* 57. ISSN: 0025-7079. URL: https://journals.lww.com/lww-medicalcare/Fulltext/2019/05001/Turning_Feed_forward_and_Feedback_Processes_on.7.aspx.

Parnas, David Lorge and Paul C. Clements (1986). "A Rational Design Process: How and Why to Fake It." In: *IEEE Transactions on Software Engineering* SE-12.2, pp. 251–257. ISSN: 00985589. DOI: 10.1109/TSE.1986.6312940.

Peffers, Ken et al. (2007). "A Design Science Research Methodology for Information Systems Research." In: *Journal of Management Information Systems* 24.3, pp. 45–77. ISSN: 0742-1222. DOI: 10.2753/MIS0742-1222240302. URL: https://www.tandfonline.com/action/journalInformation?journalCode=mmis20.

Pollack, Ari H. et al. (July 2019). "Creating synthetic patient data to support the design and evaluation of novel health information technology." In: *Journal of Biomedical Informatics* 95, p. 103201. ISSN: 1532-0464. DOI: 10.1016/J.JBI.2019.103201.

Preece, Jenny et al. (2015). "Interaction Design: Beyond Human-Computer Interaction CLICK HERE FOR DOWNLOAD." In: pp. 385–387. URL: https://books.google.com/books/about/Interaction_Design.html?hl=no&id=n0h9CAAAQBAJ.

Pye, A. D. et al. (June 2009). "A review of dental implants and infection." In: *Journal of Hospital Infection* 72.2, pp. 104–110. ISSN: 0195-6701. DOI: 10.1016/J.JHIN.2009.02.010.

"Quality registries." URL: https://skr.se/en/kvalitetsregister/omnationellakvalitetsregister.52218.html. *Last retrieved 19.08.2021.*

R Core Team. "R: A Language and Environment for Statistical Computing." URL: https://www.R-project.org/. *Last retrieved 23.08.2021.*

Raikar, Sonal et al. (Nov. 2017). "Factors Affecting the Survival Rate of Dental Implants: A Retrospective Study." In: *Journal of International Society of Preventive & Community Dentistry* 7.6, p. 351. DOI: 10.4103/JISPCD.JISPCD{_}380{_}17. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5774056/.

Randell, Rebecca et al. (2019). "Requirements for a quality dashboard: Lessons from National Clinical Audits." In: *AMIA ... Annual Symposium proceedings. AMIA Symposium*, pp. 735–744. ISSN: 1942-597X. URL: https://pubmed.ncbi.nlm.nih.gov/32308869%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7153077/.

"Reactivity - An overview." URL: https://shiny.rstudio.com/articles/reactivity-overview.html. *Last retrieved: 25.11.21.*

Reiter, Jerome P. et al. (June 2014). "Bayesian Estimation of Disclosure Risks for Multiply Imputed, Synthetic Data." In: *Journal of Privacy and Confidentiality* 6.1. DOI: 10.29012/jpc.v6i1.635. URL: https://journalprivacyconfidentiality.org/index.php/jpc/article/view/635.

Reszel, Jessica et al. (2019). "Use of a maternal newborn audit and feedback system in Ontario: a collective case study." In: *BMJ Quality & Safety* 28.8, pp. 635–644. ISSN: 2044-5415. DOI: 10.1136/bmjqs-2018-008354. URL: https://qualitysafety.bmj.com/content/28/8/635.

"rmarkdown: Dynamic Documents for R." URL: `https://github.com/rstudio/rmarkdown`. *Last retrieved: 16.12.21.*

RStudio Support. "Scaling and Performance - Tuning Applications in Shiny Server Pro." URL: `https://support.rstudio.com/hc/en-us/articles/220546267-Scaling-and-Performance-Tuning-Applications-in-Shiny-Server-Pro`. *Last retrieved: 05.01.22.*

Santos, Marco and Henrik Eriksson (2014). "Making quality registers supporting improvements: a systematic review of the data visualization in 5 quality registries." In: *Quality Management in Healthcare* 23.2, pp. 119–128. ISSN: 1063-8628.

Sarikaya, Alper et al. (2019). "What Do We Talk About When We Talk About Dashboards?" In: *IEEE Transactions on Visualization and Computer Graphics* 25.1, pp. 682–692. ISSN: 1941-0506. DOI: `10.1109/TVCG.2018.2864903`.

Seaman, Carolyn B. (1999). "Qualitative methods in empirical studies of software engineering." In: *IEEE Transactions on Software Engineering* 25.4, pp. 557–572. ISSN: 00985589. DOI: `10.1109/32.799955`.

Shekelle, Paul G. et al. (Apr. 2006). "Costs and benefits of health information technology." In: *Evidence Report/technology Assessment* 132, pp. 1–71. ISSN: 1530-4396. DOI: `10.23970/AHRQEPCERTA132`. URL: `http://europepmc.org/books/NBK37988%20https://europepmc.org/article/NBK/nbk37988`.

"shinydashboard." URL: `https://rstudio.github.io/shinydashboard/`. *Last retrieved: 23.08.2021.*

Soltana, Ghanem et al. (Nov. 2017). "Synthetic data generation for statistical testing." In: *ASE 2017 - Proceedings of the 32nd IEEE/ACM International Conference on Automated Software Engineering*, pp. 872–882. DOI: `10.1109/ASE.2017.8115698`.

Soneson, Charlotte et al. (Sept. 2020). "ExploreModelMatrix: Interactive exploration for improved understanding of design matrices and linear models in R." In: *F1000Research* 9, p. 512. DOI: `10.12688/F1000RESEARCH.24187.2`. URL: `/pmc/articles/PMC7359746.2/%20/pmc/articles/PMC7359746.2/?report=abstract%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7359746/`.

"Stadieinndeling." URL: `https://www.kvalitetsregistre.no/stadieinndeling`. *Last retrieved: 31.01.22.*

Stattin, Pär et al. (2016). "Dashboard report on performance on select quality indicators to cancer care providers." In: *Scandinavian Journal of Urology* 50.1, pp. 21–28. DOI: `10.3109/21681805.2015.1063083`. URL: `https://doi.org/10.3109/21681805.2015.1063083`.

Tay, Yangshun. "In-Depth Overview." URL: `https://facebook.github.io/flux/docs/in-depth-overview/`.

The Standish Group. "The Chaos Report." URL: `https://www.standishgroup.com/sample_research_files/chaos_report_1994.pdf`. *Retrieved on: 27.04.22.*

Therneau, Terry M. (2022). "A Package for Survival Analysis in R." In: URL: `http://cran.r-project.org/package=survival`.

Tufte, Edward (2001). *The visual display of quantitative information.* 2nd ed. Cheshire: CT: Graphics Press, pp. 96–96.

"Use containers to Build, Share and Run your applications." URL: `https://www.docker.com/resources/what-container`. *Last retrieved: 23.11.21.*

Vågenes, Lars (2022). "Developing a Data Model and an Architecture for a Dental Implant Quality Registry." In: *Master thesis in Software Engineering, Department of Informatics, University of Bergen, Norway.*

Walonoski, Jason et al. (Mar. 2018). "Synthea: An approach, method, and software mechanism for generating synthetic patients and the synthetic electronic health care record." In: *Journal of the American Medical Informatics Association* 25.3, pp. 230–238. ISSN: 1067-5027. DOI: 10.1093/JAMIA/OCX079. URL: https://academic.oup.com/jamia/article/25/3/230/4098271.

Weiss, Deborah et al. (2018). "Effect of a population-level performance dashboard intervention on maternal-newborn outcomes: an interrupted time series study." In: *BMJ Quality & Safety* 27.6, pp. 425–436. ISSN: 2044-5415. DOI: 10.1136/bmjqs-2017-007361. URL: https://qualitysafety.bmj.com/content/27/6/425.

West, Vivian L et al. (Aug. 2014). "Innovative information visualization of electronic health record data: a systematic review." In: *Journal of the American Medical Informatics Association* 22.2, pp. 330–339. ISSN: 1067-5027. DOI: 10.1136/amiajnl-2014-002955. URL: https://doi.org/10.1136/amiajnl-2014-002955.

Whiting, Mark A et al. (2008). "Creating Realistic, Scenario-Based Synthetic Data for Test and Evaluation of Information Analytics Software." In: *Proceedings of the 2008 conference on BEyond time and errors novel evaLuation methods for Information Visualization - BELIV '08*. DOI: 10.1145/1377966.

Wickham, Hadley and Garrett Grolemund (2017). *R for Data Science: Import, Tidy, Transform, Visualize, and Model Data*. 1st. O'Reilly Media, Inc. ISBN: 1491910399.

Wright, Melanie C et al. (2019). "Critical care information display approaches and design frameworks: A systematic review and meta-analysis." In: *Journal of Biomedical Informatics* 100, p. 100041. ISSN: 1532-0464. DOI: https://doi.org/10.1016/j.yjbinx.2019.100041. URL: https://www.sciencedirect.com/science/article/pii/S2590177X1930040X.

Yen, Po Yin and Suzanne Bakken (May 2012). "Review of health information technology usability study methodologies." In: *Journal of the American Medical Informatics Association : JAMIA* 19.3, p. 413. ISSN: 10675027. DOI: 10.1136/AMIAJNL-2010-000020. URL: /pmc/articles/PMC3341772/%20/pmc/articles/PMC3341772/?report=abstract%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3341772/.

Yigitbasioglu, Ogan M. and Oana Velcu (Mar. 2012). "A review of dashboards in performance management: Implications for design and research." In: *International Journal of Accounting Information Systems* 13.1, pp. 41–59. ISSN: 1467-0895. DOI: 10.1016/J.ACCINF.2011.08.002.

Yoo, Junsang et al. (Nov. 2018). "A Real-Time Autonomous Dashboard for the Emergency Department: 5-Year Case Study." In: *JMIR mHealth and uHealth* 6.11. ISSN: 22915222. DOI: 10.2196/10666. URL: /pmc/articles/PMC6284143/%20/pmc/articles/PMC6284143/?report=abstract%20https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6284143/.

Zerdick, Thomas. "Is the future of privacy synthetic? — European Data Protection Supervisor." URL: https://edps.europa.eu/press-publications/press-news/blog/future-privacy-synthetic_en. *Last retrieved: 13.01.22.*

Zhou, Yi et al. (2020). "MEPHAS: an interactive graphical user interface for medical and pharmaceutical statistical analysis with R and Shiny." In: *Bioin-*

formatics 21.183. DOI: 10.1186/s12859-020-3494-x. URL: https://doi.org/10.1186/s12859-020-3494-x.

Zhuang, Mengdie et al. (2020). "A Framework for Evaluating Dashboards in Healthcare." In: *arXiv preprint arXiv:2009.04792*.