TALLINN UNIVERSITY OF TECHNOLOGY
School of Information Technologies

German Ivanov 192929IVSB

A Secure Infrastructure Solution for a Plastic Surgery Clinic

Bachelor Thesis

Supervisor
Kaido Kikkas
Doctor of Philosophy (PhD) in Engineering

Tallinn 2022

TALLINNA TEHNIKAÜLIKOOL
Infotehnoloogia teaduskond

German Ivanov 192929IVSB

Turvaline taristulahendus plastilise kirurgia kliinikule

Bakalaureusetöö

Juhendaja
Kaido Kikkas
Tehnikateaduste doktor

Tallinn 2022

Author’s declaration of originality

I hereby certify that I am the sole author of this thesis. All the used materials, references to the literature and the work of others have been referred to. This thesis has not been presented for examination anywhere else.

Author: German Ivanov ....

(signature)

Date: April 24, 2022

Abstract

This thesis aims to examine the infrastructure of a plastic surgery clinic. The main issue addressed by this work is the security issues of small businesses and healthcare organizations in general which are targeted more frequently in recent years. This work will improve the situation for a specific organization. The research will include an analysis of a secure infrastructure as a concept and how it is achieved. The outcome of this work is an improved version of infrastructure based on the findings.

The result is a ready-to-be-deployed infrastructure aimed at handling a small load suitable for the healthcare organization in question. It uses Yandex.Cloud as a cloud platform for hosting and data storage. An open-source software OpenMRS has been used as a platform for healthcare data management. It also includes backup policies and general secure behaviour guidelines.

The infrastructure is already functioning inside the organization. The reception of it is good because it has increased the productivity of regular employees. The new deployment is used concurrently with the old approach. It makes the gradual migration of data possible. The current plan is to completely phase out the old deployment after a half year in case there are no issues with the new deployment. The old deployment will be preserved as an encrypted backup.

The thesis is in English and contains 24 pages of text, 8 chapters, 1 figure.

List of abbreviations and terms

2FA	Two-factor authentication
AWS	Amazon Web Services
GCP	Google Cloud Platform
GDPR	General Data Protection Regulation
IAM	Identity and Access Management
IaC	Infrastructure as code
VPN	Virtual private network

Table of Contents

List of Figures

1. Introduction

The main problem this thesis tries to address is that security is usually neglected in small scale businesses. There are multiple reasons which are out of the power of their owners - lack of finances and personal understanding of priority [1]. This might backlash and cause some damage and loss to those businesses. These cases are not widely discussed and not noticed unless a major hacking attack happens which affects many companies at once This has both direct and indirect effects. Businesses lose the reputation and trust of their customers which in turn decreases monetary turnover and makes it harder to continue existing as a functioning entity.

However, even though the damage is already considerable for generic small businesses. It is especially problematic for companies involved in the healthcare industry [2]. The data is more sensitive in their context and thus the cost of a potential data breach is bigger and could ruin the business more. This creates a dangerous situation where on one hand those clinics possess a highly valuable asset while being a relatively easy target for criminals to attempt to be hacked. This leads to a situation when in recent years healthcare organizations are becoming more and more frequently a target of those attacks which causes a lot of damage to both businesses and individuals [3]. This could be solved by applying modern approaches of DevOps and infrastructure development which are widely known and used in the industry of commercial software development.

This thesis tries to develop an infrastructure that could help specific plastic surgery clinic improve their current situation using modern ideas of secure infrastructure and state of the art technologies. As a result, it will improve the overall stability and integrity of stored data and guarantee suitable privacy for their clients and the information they are providing them.

The chapter 2 describes the picked methodology and how the stated problem will be solved. The chapter 3 gives some background about the current state of affairs in the healthcare organization in question and what are the different unique specifics of it. The chapter 4 describes what is a secure infrastructure and what has to be done to achieve it. Besides, it introduces some of the unique security challenges of the healthcare industry. The chapter 5 examines what are the possibilities for hosting applications and what are the advantages and disadvantages of them. The chapter 6 explains the decisions behind picked tools for the development of mentioned infrastructure. The chapter 7 contains a detailed description of the developed infrastructure and what are the plans related to its deployment of it.

2. Methodology

This thesis will be action research. It will attempt to solve a real-world issue and study the experience of this process. The goal is to improve the observed situation and possibly some of the outlined issues. The quality of an outcome can be assessed based on the knowledge outcome of participants of this work [4].

Investigation of existing works regarding common security pitfalls in healthcare

To be able to come up with good infrastructure, it is important to understand what are the common or specific security risks in the healthcare industry. The situation is different between industries so it is important to research that beforehand. This will allow finding a reasonable balance between complexity and cost.

Secure Infrastructure overview

It is important to understand the concepts and key values behind secure infrastructures. To achieve it, it is needed to research the current state of the art when it comes to the development of such infrastructures. This gives some key aspects which should be considered during this thesis and in turn, will make it possible to deliver a better solution.

Hosting platforms comparison

The platform is the foundation of any infrastructure. There are multiple options to choose from. Mainly major cloud providers and private service providers. Both approaches are commonly used in the industry. However, they have different security, complexity and cost implications. The choice might also depend on the specific requirement (e.g. compliance with local regulators). The budget possibilities might also influence the decision.

Choice of tools

The small size of a business implies the infrastructure should be relatively simple to operate and manage. To achieve it, it is important to use well-known and widely used tools for the management of such infrastructure. This will ensure the maintainability of such a project which will make it easier in the future to attract contractors to iterate on the current setup if the need arises.

Configuration development

The ready-to-deploy configuration for the infrastructure will be stored as a git project. In addition, the decisions and trade-offs made during this development will be documented and explained. The picked storage will make it easy for anyone interested to hand it over to new maintainers or use it for their own purposes.

Result overview

The developed infrastructure is implemented and the outcome of it is reviewed. The basis of such review is the principles of secure infrastructure which were outlined in previous chapters. The knowledge outcome and positive impact in regards to the original infrastructure have to be evaluated as well.

3. Background

The plastic surgery clinic in question is a small scale medical organization based in Moscow, Russia. It has approximately 10 employees at the moment which includes the management and medical staff. It has experienced several outages which influenced its work and made them suffer some financial losses. They decided not to invest in an upgrade because they considered it to be a financial burden without clear benefits. They have agreed to participate in this research on the grounds of anonymity and agreed to try out the proposed solution in case they decide it is worth it.

They have 10 employees at the moment. 7 of them are management and 3 of them are medical staff. They all have full access to internal databases and management systems. There are no access restrictions based on user roles or anything similar to that. The used software is also available for everyone. It is not restricted to an internal private network so anyone with internet access has a chance to try to log into their systems. Every employee has a separate account with unique login details.

The current infrastructure is only one server which is located in the building of this clinic. This server is fully owned by the clinic. In theory, it is possible to get direct physical access to this server for every employee. The server is maintained by an external contractor on-demand basis. In general, there is no one responsible for the active monitoring of that server. The contractor is called in case any system failures are experienced. It was noted these services are rather expensive for the clinic.

The clinic uses a mix of paper-based and computed based methods of data management. The computer-based method is a custom web application written for this specific clinic. It was written once by the contractor and was never extended nor maintained since then. The app is not customizable so to satisfy the need for new ways of data management, it was decided to use a paper-based approach to close that gap. The clinic has expressed a desire to unify this process.

The clinic has never had any governmental security audits. Due to the specific local regulations, clinics of that size are not required to have any certificates. This simplifies the requirements to fit during platform and software changes. It is not as strict and leaves a significant amount of flexibility. The clinic is also not collaborating with other healthcare organizations so there is no need for some common ways of exchanging data between businesses.

4. Security

This chapter gives a theoretical background of security infrastructure and the main security risks of healthcare infrastructure.

4.1 Healthcare industry problems

Malware is one of the most popular tools to abuse the vulnerabilities of insecure infrastructures. [5] That kind of ransomware usually utilizes two main aspects: outdated software and poor access control. The former is a source of already known security breaches in the used software which is later abused to get direct control over different parts of the system and as a result of the consumer data. The latter makes it easier for criminals to infect all devices in the network. What is even worse, depending on the setup in a particular clinic, there might be no need to infect all devices in a network, if every device of such network has full access [6].

Another specific of the healthcare industry is the special medical equipment widely used in hospitals and small scale clinics. They used to be overlooked and ignored during safety audits based on the assumption of safety. As it turns out, such an assumption is not entirely correct [7]. There is a special type of attack called "Medjack". It is an exploit that spread its malware using unprotected medical equipment [6]. During a relatively short period, this kind of exploit can infect all devices from the inside out. Resulting in dangerous consequences. Such malware can fully block healthcare organizations from operating which can lead to both financial risks and risks to human lives.

Different security researchers have noted that the equipment is a "security nightmare" [8]. Even though such attacks were not recorded in the wild, the chances of such attacks happening will only grow in the following years due to the potential gain for criminals [9]. It can be used to cause hardware to malfunction resulting in a potential death of a patient due to the problems during operation, fatal overdose or stopped life stopped devices [10].

This must be addressed both on the organizational as well as on the manufacturer level. Organisations should strive to build a secure infrastructure that would minimize the result of a potential hacker attack and thus decrease potential damage. Besides, healthcare organisations are advised to organise extensive internal training in regards to digital sanity and how every individual should behave to avoid possible phishing and other types of attacks using social engineering [11].

4.2 Secure Infrastructure

The value of a secure infrastructure is high because it prevents damage which might happen due to potential attacks. The modern IT system is usually complex and includes multiple devices of different sizes and security levels. All of them have different producers and thus different practices used during the creation. Some are highly secure and frequently updated, and some are not. Nevertheless, infrastructure is the core of any business operating in the modern world. A plan how to secure infrastructure is a master plan affecting all parts of the organization. Besides, as was discussed in the previous chapter, the healthcare industry is facing record attention from cyber-criminals [12]. Unfortunately, it is also an industry with outdated security practices and lots of equipment responsible for health yet being so vulnerable to external manipulations [12]. This created a situation where one of the most integral parts of social service is one the least protected. This increases the value of a secure infrastructure even more.

Infrastructure security is a multi-field discipline. It includes different measures to protect software and hardware assets and make sure they are safe. This includes both physical and cyber safety. It affects all devices which are interconnected and communicate with each other in a typical scenario. For example, employee notebooks, servers, phones connected to an internal Wi-Fi network, medical equipment, and networking systems - all are considered during protection activities.

Besides, infrastructure security considers physical threats in combination with hacking attacks [13]. Different natural disasters such as earthquakes and hurricanes can cause some serious complications for the outlined devices. In other words, the end goal is to provide resilience to possible threats. Ideally, a combination of measures must minimize the amount of potential downtime and as a result, minimize the number of financial losses and custom perception.

It is important to develop not only specific infrastructure but also some behavioural guidelines which would decrease the overall potentially surface. Users that are aware of dangerous behaviour are making fewer amount mistakes and leaving less space for hackers [14]. This is all done to achieve security according to CIA triads which is the general and widely regarded security model. It stands for **confidentiality**, **integrity** and **availability** [15].

A secure infrastructure must preserve confidentiality. Sensitive data must be not available to unauthorized parties. This is the case with the majority of data in healthcare. It must also ensure the integrity of the data. It must try to keep data consistent and modified by authorized users. It is also rather important for clinics and hospitals. Maliciously modified data can cause lethal cases for the client. For example, the data about some allergies might be deleted from the system which would result in the administration of drugs incompatible with the life of a patient. Availability is a no less crucial aspect for organizations involved in the health maintenance of users. It is more strict in big hospitals but is nevertheless important for small scale operations. On one hand, it is intolerable to have any downtime during surgeries, on the other hand, it might affect the financial indicators of a business.

Infrastructure security involves multiple layers going from the most high level starting with physical security and ending with data security. The development of a secure infrastructure must output a solution which would address security on all of those levels. Whether it is an advanced data centre addressing the physical security or thought through backup policies addressing the data security.

The general approach is defining 4 levels of security: physical level, network level, application level, and data level. The clinic in question has control over the majority of these levels so it can vastly improve its current security. It only cannot influence the application level. It usually implies hardening the developed application against different attacks. The clinic is planning to use already existing applications for the management of its data so the only possible action in that realm is to pick the most supported software.

A physical level included all real-world aspects of the IT system. All servers are located somewhere and powered somehow. It is important to protect serves from natural hazards and direct access to hackers. A network level is how information is travelling between different parts of the system. For example, this includes a firewall or separation of the internal network from the internet using a VPN. The application-level was already previously discussed. The data level is about protecting data in general. This includes encryption, backups and anonymization. All of that is to preserve data in case of any problematic situations.

Such a holistic approach to security allows to build a solid framework that addressed potential issues on multiple levels. In general, it is easier to address complex issues using decomposition. Big problems are split into smaller, manageable pieces. The 4 level approach tries to achieve exactly that. It makes it possible to address different sets of attack vectors separately and then combine them into one.

Many measures can be used to increase the safety of a solution. It will be discussed some of the measures that in combination create a secure infrastructure: physical hardware safety, IAM (Login system) and secure data storage practices. [16]

Physical hardware safety is the first layer that must be considered during the evaluation. Different data centres have different levels of access security, some organizations are managing their servers privately. There are even cases when servers are not in any specific data centre but rather in some general business building. This is a very dangerous situation. If this link of a chain is weak, then everything else does not matter. There are multiple risks associated with the usage of questionable security practices when it comes to hardware safety. Criminals can easily abuse it and get direct access to servers running code that will allow them to perform any actions. As a result, this will completely mitigate all digital security measures and make them useless [17]. Physical access allows them to bypass them all. Besides, another important factor that must be considered and is usually handled by data centres is natural hazards. This is something that might significantly influence the data integrity and stability of the hardware [13].

Advanced IAM is another good security practice for any business managing sensitive data. It must be a multi-level system with fine-grained control over the users and their roles. Every involved user must have the least possible amount of reading and writing rights. There are multiple reasons for such restrictive access. The more restrictive such system is, the better for the organization it will be. As a big business entity, it is hard or almost impossible to actively track every employee in every office and how secure and honest their behaviour is. It also means the business has to trust its employees which are already at risk. However, if a highly restrictive IAM system is in place, then such risk is smaller. In case of a trust breach from the employee side, the damage will be minimized because every single employee usually has a very small scope in the scale of the whole business unit. An additional useful tool that is provided by a good IAM system is a detailed logging system. Such a system created a log of events inside a company. It can give transparency into all actions that are performed by employees: reads, writes, deletes. Every data change and action inside the system must be possible to retroactively investigate. This will help with the research of possible trust breaches as well [18].

Sophisticated login system plays a role in the insurance of minimal risk of access keys leak. The more personalized it is, the harder it is to steal. The login system must be created in such a way that the human factor would be minimized as much as possible. People tend to use simple passwords, which the system should not allow them to do. In the same fashion, the system should not allow them to use passwords that are present in publicly available databases of leaked passwords.

Two-factor authentication must be enabled by default for accounts as well. This will help in case a person gets phished and as a result, leaks his login details [19]. These tools can help with these cases. If the 2FA is enabled, then login details will not be useful because the 2FA is unique for the login attempt. The same applies to biometric-based authentication protocols. They are the most optimal because they are the hardest to steal without the direct collaboration of an employee with criminals [20].

Data Anonymization can be practised in the accordance with GDPR. This implies good data handling practices. For example, it means that the company must collect a minimal amount of data. The less company collects data, the less it can leak in case of a data breach. To secure data furthermore, it is advised to use different anonymization methods on that data. In an ideal case, the stored data should be impossible to be linked back to the original owner without significant time costs. This decreases the potential damage of a successful attack and in some cases might make it unfeasible to carry out. This also works not only for the criminals outside the organization but also for wrongdoers inside the organization. Some people will have direct access to customers' data as a part of their daily work, this means it is a potential point of failure and leak. However, if the data is anonymized, it will not be possible to use that against users.[21]

Backups are another important part of a secure infrastructure. It ensures that important data will be easily recoverable in case of any system failures. This is also important to make those backups automatically because only this will be a consistent and reliable way of making those backups. A good practice is to clone them between multiple independent systems [22]. If backups and the live data are located on the same machine, then the value of such backup is highly questionable. The reason is that in the case of a system failure, the backup will be lost as well so there was no reason to create it in the first place.

Regularly, backups must be tested against the existing documentation of data recovery. This will ensure that those backups will be possible to use and that employees are aware of the needed steps to restore them. In case it is not done, in a critical moment, it might be discovered that backups are either corrupted or recovery instructions are not clear enough. For example, there is a concept of "Chaos Day" when some parts of the infrastructure are destroyed on purpose to see how the team can react to it. [23]

5. Platform

There are two competing ways of hosting your platform using a cloud or choosing on-premise solutions. The choice will greatly depend on the context. There are different situations when one option might be more favourable than the other. In the context of a small scale healthcare organization, it means there are multiple constraints: a limited budget and the requirement of outsourcing as many parts of the infra as possible due to the interest of minimizing expenses.

It is an ongoing discussion on which choice is better for different cases. The discussion is only about the financial part of the question but also about the influence the cloud computing movement has on the tech industry as a whole. There are some concerns about the growing centralization of power in hands of big based USA tech giants [24]. There were already some cases of power abuses by those companies [25]. Considering the recent licence change of an Elastic search, this would harm them financially Amazon so they have decided to fork it and continue maintaining their own version [26]. This is only possible because Amazon has grown so big that they can allow themselves to behave in this way.

The political involvement of those companies is another factor that might influence the choice. Especially in the context of a company working in Russia. The country has some difficulties in diplomatic relationships with the US [27]. There were cases where some companies were cut off from those cloud providers due to the disagreement in political views with the owners of the platform [25]. This makes cloud owners rather vocal and not independent in their service providence. This might be a cause of some potential unexpected cancellation of service.

Also, there are some legal requirements to satisfy. For example, in Russia it is obligatory to store all user data on servers inside the country. [28]

Considering these specifics, it is important to analyze both solutions and see how they perform in different aspects of infrastructure development which are considered important.

5.1 Security

In general, cloud providers are considered more secure, especially for small-sized companies [29]. It is technically possible to build a private data centre of the same grade but it will be possible only for companies of significant size. Besides, there are some products provided by those cloud providers which can significantly improve security. As an example, 2FA or VPN solutions are always available for major cloud providers. Building those solutions from a scratch might be cost and hardly imaginable for small companies. Those tools are also frequently audited which increases the safety of those solutions. This is usually lacking for custom solutions. However, if one decides to build them from the ground, it usually implies the employment of many high-cost professionals or security firms, which might add significant time and money expenses. This will also make future management of such a platform more complicated because it won't be well known by other potential employees.

In conclusion, it is possible to say that the choice will significantly depend on the size of a company. It seems that for small-sized companies more appropriate would be to choose existing cloud providers which would give them reasonable security for the amount of money they are paying. Big companies will have flexibility in that question. They have at least a theoretical capability to gain the same amount of security in their private data centre if they decide to build it.

5.2 Disaster recovery

Another important factor to consider is the capabilities of disaster recovery. This is where high-grade data centres might be more beneficial in comparison with in-house hosting. When things are going well, they might seem pretty similar in capabilities. However, when a power outage happens or there is a hurricane happening in the area, the difference will be significant. This is important to consider during hosting cost calculations. That kind of outcome is rare and low-risk. However, the impact of such a disaster might be significant.

Cloud providers are providing solid disaster recovery [30]. In a typical on-premise solution, the need for a multi-region deployment across different data centres to provide required data redundancy might be very expensive and labour heavy. In contrast, cloud providers are providing such solutions for a smaller price because the required infrastructure is existing and there is no additional cost to provide it to new customers. They are also controlling a significant amount of computing power which allows them to negotiate better prices and get better terms, to satisfy other needs. For example, GCP was able to switch fully to green energy because they are building data centres themselves [31]. This allows small businesses to experience the same benefits, some other bigger companies might experience, for a smaller price.

5.3 Elasticity

Cloud providers are owning a significant amount of computing power. They have lots of data centres in all parts of the world. Consider GCP which is owned by Google, there has a data centre in almost every western European country [32]. All of that computing power can be provided on demand to their customers and it gives results another important advantage of cloud computing - elasticity.

In other words, it is the possibility to quickly increase the amount of available computing power to handle either increased load due to the user activity or increased data storage requirements. This is usually rather expensive and complicated because required physical changes to the setup or software changes to handle that change. This also might bring some financial benefits. Some cloud providers might offer lower prices in you are using a bigger amount of servers from them [33]. However, this advantage mostly applies to services that cannot estimate the user load or are experiencing a sudden increase in the load at times. This doesn’t apply to small plastic surgery clinics where the traffic and load are very well predicted. However, it is possible that the clinic will continue developing and open some similar clinics in other parts of the country. In this case, the elasticity and possibility to easily meet any computing demand will be handy.

5.4 Infrastructure outsourcing

The cost of cloud computing is complex and case-specific. There are situations that might be more expensive or not. It is important to consider the specifics of your situation. However, as a general rule of thumb, it is possible to say that for small companies, the cloud can save quite a bit of money [34].

Cloud providers are cheaper in the long run due to the decreased maintenance cost and the absent need to hire additional infrastructure people. This is especially important for small companies that cannot afford professionals who could provide high-quality service. Also, the knowledge of specific cloud providers is transferable and thus the supply of relevant specialists is higher and easier to find. On top of that, the total cost of ownership is usually smaller because there is no need to maintain hardware, organize the maintenance of that hardware or invest in the outage prevention system, and hardware upgrades.

Besides, if you are using managed products of those cloud providers. For example, managed DB instances. It gives you the expertise of employees working for Amazon, Google, IBM or Microsoft. However, you are not paying more for that and you do not have to spend time on hiring such high-grade professionals. This is possible thanks to the operational modal clouds have decided to use.

5.5 Community support

There is also the re-usability factor of clouds. They are providing many premade solutions for different needs starting from the user management and ending with the advanced IAM. This can significantly help with getting high-quality solutions with verified security for a much lower cost. This is especially important in cases where is no need to develop a novel approach but there is a need to use already existing industry-grade solutions which is the case for the healthcare industry.

It also makes it easier to develop infrastructure on the cloud because there is a lot of open-source knowledge and shared configurations that can be used due to the flexibility of a cloud. This can decrease time-to-market significantly and as a result, start bringing value in a much shorter time frame than if everything was developed from scratch. There are also many volunteers who are willing to help in case you are in doubt.

5.6 Self-service

Cloud providers are investing a lot into cost optimizations. One of the possible cost optimizations is automatizing as many aspects of the infrastructure as possible. The more client can do on their own, the less they will reach out to customer support or need any additional assistance. This is useful not only for the cloud but also for customers as well. This simplifies already existing procedures.

Usually, internal tooling is a good way of increasing productivity and decreasing the number of errors. Cloud providers are providing many self-service tools which can help to automate and speed up many processes related to infra management and provisioning. Besides, the community around those cloud providers have built many tools which can further benefit the business.

6. Choice of tools

In the context of a small plastic surgery clinic, the optimal approach is to optimize the usage of already existing solutions and minimize any custom software development. The budget for infrastructure is limited. Besides, the clinic is not required to gain any security certification because it is not required in the country it operates. This makes the pool of choices wider than it could be.

6.1 Platform

Based on the research, the cloud platform is more appropriate for the specific client. It gives the benefits of required data security, integrity and a wide selection of already developed products that clients might make use of. This choice will also allow to automate infrastructure-related operations and speed up the development. Also, this will simplify the execution of security best practices such as automatic backups, advanced IAM and multi-region setup.

The choice of a specific cloud is more complex. There are multiple major cloud providers: GCP, AWS, Azure and other smaller providers. They all share common cloud computing advantages and disadvantages. There is ongoing active competition among these industry leaders. So, it is hard to pinpoint the clear best choice for all cases. Usually, the choice depends on the experience of a specific developer owning the development of an infrastructure solution or the need for a niche product provided only by a given company.

It must be noted that, in the industry, it is regarded as a good practice to develop infrastructure without relying on some unique characteristics of a cloud provider to avoid so-called "vendor lock-in" [35]. This will significantly decrease the complexity of a potential cloud provider switch in the future.

For a given company, there are none of such special needs for some specific software so it does not influence the final decision. However, some unique requirements must be satisfied. The cloud provider must have servers located in Russia because the law requires it [28]. Besides, there are recorded cases when governmental organizations have blocked wide ranges of IPs of different cloud providers which complicated the access to them

[36]. This could significantly influence the availability of a deployed system. As a result, it is a necessity to use a local cloud provider. The main players in the market: Yandex, Mail.ru, SberCloud. The Yandex.Cloud has the biggest market share [37]. It is also backed by Yandex which is a major player in the Russian market. This might simplify future maintenance and decrease the chances of unexpected bankruptcy of the hosting company. It was decided to use Yandex.Cloud as a platform provider for this project.

6.2 Configuration management

6.2.1 IaC

Infrastructure configuration management is a complex and error-prone matter. The most modern approach to addressing multiple common complexities is a concept called Infrastructure as Code (IaC). IaC's main idea is the requirement to describe your infrastructure as a set of pieces of code that are stored in an organized manner. That code can be later automatically executed and it brings those life changes [38].

This approach has multiple benefits compared to the old way of doing things when configuration on was done on case-by-case basins directly on already working parts of the infrastructure. First of all, it significantly simplifies the disaster recovery in case any of the services get corrupted and lost their existing state because recovery of such service will require execution of a few simple commands. Secondly, it brings more transparency to the infrastructure of a company since everything is now located in the same place [39].

What is important in a stability critical industry such as health care, IaC reduces the number of potential bugs. It eliminates the risk of configuration drift between the one running on services and the one developer is editing. Besides, it is possible to use this approach with already existing control version systems such as git. This makes the investigation of any configuration changes effortless and streamlined [39]. Besides, IaC tools are usually not tied to a specific platform so it will decrease the factor of a vendor and simplify possible future migrations. IaC is also easier to document using existing developing tools which simplify project handover to future developers.

Many tools can help with the IaC execution. Both vendor dependent and not. The leading independent open-source ones are usually named Ansible, Chef, Terraform, etc [40]. All of these tools have their benefits and disadvantages. Among these, the most optimal choice is Ansible. It strikes a great balance between simplifying and offering functions. It is very well known in the industry and has existed for a long time (10 years[41]). It is properly supported by all cloud vendors and has all the needed functionality for a modern IaC. It is also backed by a big corporation, Red Hat, which makes it a safer choice with the low choice of the sudden drop in maintenance which is an undesirable outcome for a healthcare company. It is agent-less which makes it easier to set up and manage than other competing solutions. Overall, it is a solid pick as the main tool for IaC because it is safe, tested, well-known and will not complicate the search for new developers.

To simplify the development, we will utilize as many Yandex.Cloud products as possible because are highly compatible between each other, thought-through and, as a result, there is no need to reinvent a wheel. However, some specific healthcare tools are not provided by Yandex.Cloud. For example, the healthcare management system is a core of any clinic and must be present. The already existing custom PHP application is decided to be replaced by some more widely used in the industry alternatives. The downsides of an existing solution are the maintenance complexity and the absence of any security audits. Both of these are hardly fixable due to the financial difficulties of the client. The most optimal choice would be to find a stable open-source application that would strike a reasonable balance between the security and price.

6.2.2 Yandex.Cloud

DB must use managed solution because data is the most critical part of this organization. This will ensure high availability and easy backup mechanics, and recovery. **Yandex Managed Service for MySQL** fits the outlined requirements. They provide by backups and legally guarantee data durability. They will also keep up the software updated and apply the latest security patches to eliminate the risk of so-called 0-day exploits. They support a wide range of DB types so they can be used with the majority of applications.

File storage will use Yandex Object Storage which is a cost-effective data storage solution [42]. This will ensure backups do not cause any financial issues for the client.

Yandex Virtual Private Cloud is a product that will be used to protect the network part of the infrastructure. It will be used to ensure that only authorized users with the right intent will be to reach our infrastructure. And in case any wrongdoers do cause some harm, they will be stopped from doing so.

As an app host, it is possible to use **Yandex Compute Cloud** which is a product that encapsulated the management of custom virtual private which can be customized to run any software depending on the need. They are scaleable so in case there will be increased demand for computing power, it will be easy to satisfy.

6.2.3 Open Source Solutions

As an internal healthcare management system, **OpenMRS** will be used. It was initially developed for usage in developing countries where budget and computing power are very limited [43]. It is highly versatile and is supported by a wide-open source community. It was developed to be extensible and to be easily adaptable for different use cases. One of the great features is the possibility to add new functionality without coding so it is an interesting option for small scale clinics without significant budgets available for the additional R&D [44].

It is actively developed and supported by many health organizations around the world which give confidence in the future of this project [45]. The source code is distributed for free and is accessible to anyone interested. So, it is not closed for security audits. This has been already used for penetration testing by different individuals. As a result of these, several security issues were fixed. This wide support and financing mean the software will continue to get security patches and updates so it is a balanced choice. It is also frequently a part of the Google Summer of Code [46]. This is an event where many software developers are trying to improve already existing software projects [47]. That kind of participation ensures a consistent yearly flow of improvements and reviews.

Since it was designed to work in hard conditions, it will not require significant computing power and as a result, will be both cheaper and easier to maintain. This makes it even a better choice for the client. This software uses MySQL database which is supported by **Yandex Managed Service for MySQL** and is written using Java which is possible to run using **Yandex Compute Cloud** [48]. It means it will be possible to deploy that piece of software using already selected platforms and products.

Since it is a flexible solution, it will be possible to use this tool as a base for the optimization of all other processes inside the clinic. There are still some paper-based operations. Paper-based processes are considered more ineffective and more error-prone than their electronic alternatives [49]. The main reason in this specific clinic why paper-based information mediums are still used is the limited capabilities of their custom healthcare management system. However, since OpenMRS is designed from the ground to be easily changeable and customised, it will possible to adapt to support all unique use cases. As a result, according to the research, computer-based management systems will allow them to work more efficiently and makes fewer mistakes [49]. Both outcomes are very valuable in the healthcare industry.

7. Infrastructure development

Parts of the configuration which can be publicly shared are stored in a git repository located at `http://github.com/tasyp/infra-thesis-2022`

7.1 Overview

As a reference point for the development, the previously made analysis of a secure infrastructure was used. The 4 levelled infrastructure abstraction is used for the description of taken measures.

Physical level security was solved solely using a trusted and widely popular cloud provider - Yandex.Cloud. Its data centres are highly compliant and reliable. They are providing all the necessary equipment for the protection of assets. Besides, they are providing the basis for a successful backup policy - they have multiple servers in different regions to increase redundancy.

Network level is secured using two tools. The internal infrastructure is protected from the outer world using a VPN. This ensures all communication with services is additionally isolated and encrypted to make it harder to sniff the traffic. The firewall is configured to deny any incoming traffic in case it is not using VPN.

Application level, as it was previously mentioned, is not in the complete control of this project. However, it was decided to use an open-source application instead of a custom one. This application is a group effort and is regularly reviewed by multiple contributors. This ensures that the code is written using principles of secure programming and risks of different exploits such as SQL Injections are lower.

Data level is addressed by multiple techniques. First of all, there is a set up of automated backup policies which are regularly created and stored in different regions in an encrypted format. Secondly, the application used for healthcare management has different settings for increased anonymization of stored data.

7.2 Schema

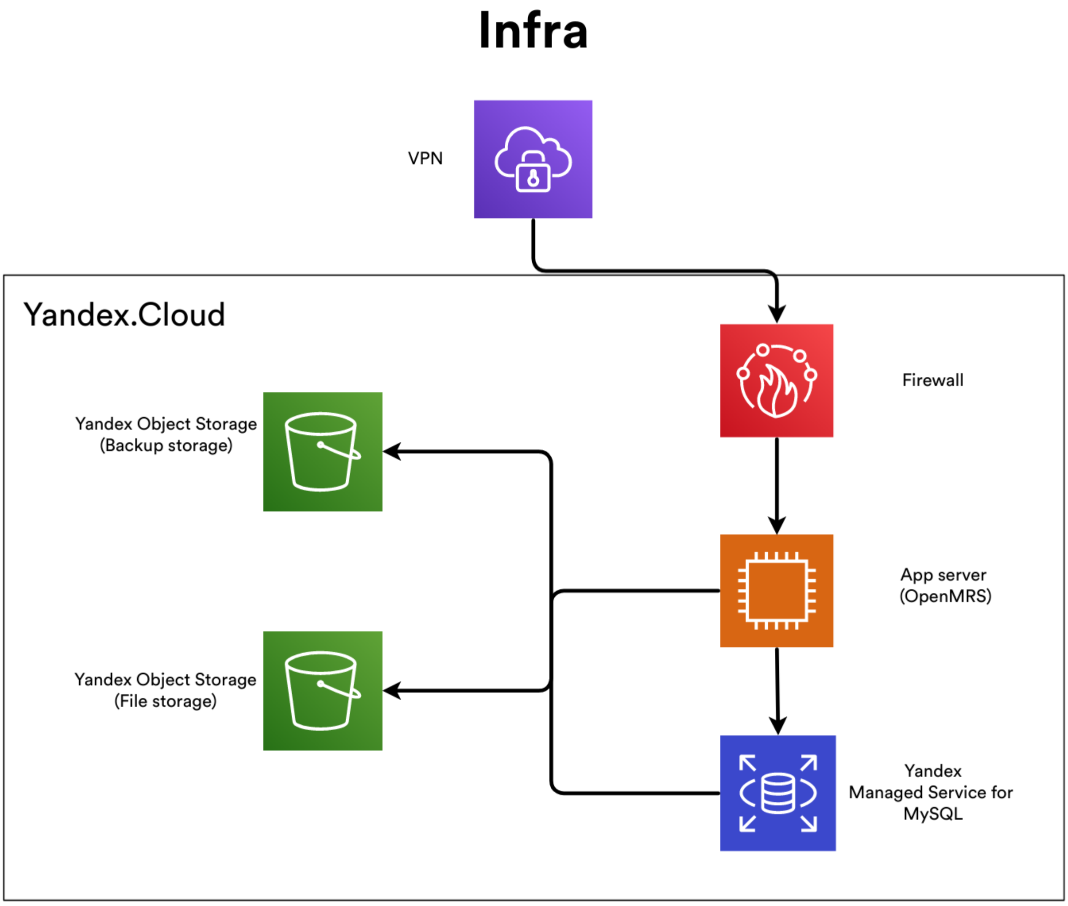

Figure 1. *Proposed infrastructure schema.*

It includes one server running app server and managed DB instance for data storage. It also includes multiple supplementary services which are used to secure this infrastructure: backup management, firewall, file storage services, vpn.

App server is a docker based container which is run on the Compute Cloud instance. Data is stored on the managed MySQL Database. It has no scaling of any kind enabled because the historically clinic has not seen any spikes traffic so it is unnecessary to enable it.The firewall is setup in the way that it is impossible to connect to this infrastructure without vpn. This ensures it is hard to cause any damages to this system from the outside.

7.3 Configuration

The configuration is an ansible playbook which is a set of configuration files where the main file defines the correct order of execution.

The overview of source:

```
- name: Provision OpenMRS deployment
  hosts: servers
  become: yes
  roles:
    - role: nginx
    - role: docker
    - role: app
    - role: users
```

Every included role has its own function:

- **users** creates needed admin accounts to manage the system.
- **app** provisions servers to host the health-management system

Ansible playbook is responsible for the deployment of an OpenMRS application. Other parts of the infra such as backups, VPN and Firewall are configured interactively using GUI of Yandex.Cloud.

7.4 Deployment

To deploy OpenMRS, the only needed step is to replace respective env variables and execute the following:

```
ansible-playbook playbook.yaml
```

This will execute all needed configuration and create it from the scratch on the Yandex.Cloud account. It also considers the current state of an infrastructure. If the configuration was applied before, it will not be applied again. This speeds up any iterative configuration updates after the initial release.

7.5 Results analysis

The described infrastructure is already deployed and is used inside the organization. However, the old setup is not completely phased out. The reason is a complicated migration path. There aren't any available tools which would do that automatically with existing data. This is also not possible with paper-based data. Both of these factors lead to the decision of using two systems concurrently. This will allow minimizing the number of upfront time investments for employees. However, the improved workflow will make it easier to convince for the complete switch later on. The security increase is definitely more important. The issue is not as visible for regular employees as improved workflow. Nevertheless, every employee has received a guideline for safe behaviour. This has definitely increased the awareness inside the organization. The general approval of changes has increased among employees. Based on the previous research, it is possible to say that employees will make a smaller amount of mistakes and it will also facilitate the security of an infrastructure [14].

The current plan is to wait for half a year and see if any major issues occur during this period. This amount of time should suffice to complete the migration. When migration is complete, the old system will not be used anyway because all user data has already been transferred to the new system works. This will allow disabling the old systems without causing any operational difficulties. One of the important criteria was to minimize the potential downtime during this project. Even though theoretically the need should not arise, the backup of the old system will be done. It might happen that there are some rarely used and to make sure they are recoverable and lost, a backup will be created.

When the migrations period completes and the old system is completely phased out, the story of secure infrastructure will not end for this specific clinic. This is actually a story with no clear end. There are multiple additional improvements that can be made. For example, it is possible to order external audits of cyber-security professionals to get an additional review of the current setup. It is also possible to send an application to get a security certificate. It is not a legal requirement for this clinic. However, this would help to show serious intent to their customers.

8. Summary

The research aimed to find a secure infrastructure solution for a plastic surgery clinic. The main criteria for an infrastructure that could be later implemented by the affected organization was financial load and maintenance load. Based on these requirements, an analysis was performed to see what are the current possibilities on the market which could meet these demands.

It was decided that the most optimal platform for such infrastructure is a cloud based one. Among the competitors the most popular and the current leader on the Russian market is Yandex.Cloud. It has the biggest market share and is a safe for such project. To maximize efficiency and minimise maintenance load on the healthcare organization, it was decided to use as many managed solution provided by Yandex.Cloud as possible. This ensures high performance and stability with no need of active participation on the client side. As a healthcare management system, OpenMRS was picked which is an open source product that was initially designed to work in conditions with unstable and low-tier computing power. It is highly extensible and does not require programming knowledge for such changes.

As a result, ready to be deployed configuration was developed with according backup policies and other related documentation. It was developed as a IaC (Infrastructure as code) solution. It means it is a set of instructions which describe how the infrastructure should looks like statelessly, meaning it is possible to recreate the whole infrastructure automatically. As a base tool for that, Ansible was picked. It is a well established configuration management open-source software. It is a tool that makes IaC approach possible and makes it in a platform independent fashion. The configuration is stored in a git repository due to the convenient history management and straightforward collaboration capabilities.

The final infrastructure was deployed and is now concurrently used inside a clinic to better understand with the old setup. It is done so that gradual migration could be formed and the new system could be adopted to the existing use-cases. The current reception is good and no problems were experienced during the rollout period. The migration period will continue until the old system is completely phased out.

Bibliography

[1] Loren F Selznick and Carolyn LaMacchia. "Cybersecurity liability: How technically savvy can we expect small business owners to be". In: *J. Bus. & Tech. L.* 13 (2017), p. 217.

[2] Clemens Scott Kruse et al. "Cybersecurity in healthcare: A systematic review of modern threats and trends". In: *Technology and Health Care* 25.1 (Feb. 2017), pp. 1–10. DOI: `10.3233/thc-161263`. URL: `https://doi.org/10.3233/thc-161263`.

[3] *KPMG Cyber Healthcare Survey*. URL: `https://advisory.kpmg.us/content/dam/kpmg-advisory/PDFs/ManagementConsulting/2015/KPMG-2015-Cyber-Healthcare-Survey.pdf` (Accessed: 10.3.2022).

[4] Steve Easterbrook et al. "Selecting empirical methods for software engineering research". In: *Guide to advanced empirical software engineering*. Springer, 2008, pp. 285–311.

[5] *Hacking Still Leading Cause of 2015 Health Data Breaches*. URL: `https://healthitsecurity.com/news/hacking-still-leading-cause-of-2015-health-data-breaches` (Accessed: 18.3.2022).

[6] Sinclair Meggitt. "MEDJACK Attacks: The Scariest Part of the Hospital". In: (2018).

[7] Patricia Williams and Andrew Woodward. "Cybersecurity vulnerabilities in medical devices: a complex environment and multifaceted problem". In: *Medical Devices: Evidence and Research* (July 2015), p. 305. DOI: `10.2147/mder.s50048`. URL: `https://doi.org/10.2147/mder.s50048`.

[8] *Medical Devices Are the Next Security Nightmare*. URL: `https://www.wired.com/2017/03/medical-devices-next-security-nightmare/` (Accessed: 1.4.2022).

[9] David C. Klonoff. "Cybersecurity for Connected Diabetes Devices". In: *Journal of Diabetes Science and Technology* 9.5 (Apr. 2015), pp. 1143–1147. DOI: `10.1177/1932296815583334`. URL: `https://doi.org/10.1177/1932296815583334`.

[10] Daniel Halperin et al. “Pacemakers and Implantable Cardiac Defibrillators: Software Radio Attacks and Zero-Power Defenses”. In: *2008 IEEE Symposium on Security and Privacy (sp 2008)*. IEEE, May 2008. DOI: `10.1109/sp.2008.31`. URL: `https://doi.org/10.1109/sp.2008.31`.

[11] Guy Martin et al. “Cybersecurity and healthcare: how safe are we?” In: *BMJ* (July 2017), j3179. DOI: `10.1136/bmj.j3179`. URL: `https://doi.org/10.1136/bmj.j3179`.

[12] Lynne M. Coventry and Dawn Beverley Branley. “Cybersecurity in healthcare: A narrative review of trends, threats and ways forward.” In: *Maturitas* 113 (2018), pp. 48–52.

[13] Abhilash Panda and Andrew Bower. “Cyber security and the disaster resilience framework”. In: *International Journal of Disaster Resilience in the Built Environment* 11.4 (Apr. 2020), pp. 507–518. DOI: `10.1108/ijdrbe-07-2019-0046`. URL: `https://doi.org/10.1108/ijdrbe-07-2019-0046`.

[14] John M Blythe. “Cyber security in the workplace: Understanding and promoting behaviour change”. In: *Proceedings of CHItaly* (2013), pp. 92–101.

[15] National Research Council. *Concepts of Information Security*. National Academies Press, Jan. 1991. DOI: `10.17226/1581`. URL: `https://doi.org/10.17226/1581`.

[16] Chris Dotson. *Practical cloud security : a guide for secure design and deployment*. Sebastopol, CA: O’Reilly Media, 2019. ISBN: 9781492037514.

[17] Ivan Blagoev. “Neglected Cybersecurity Risks in the Public Internet Hosting Service Providers”. In: *Information &amp Security: An International Journal* 47.1 (2020), pp. 62–76. DOI: `10.11610/isij.4704`. URL: `https://doi.org/10.11610/isij.4704`.

[18] Kevin Hamlen et al. “Identity management for cloud computing”. In: *Proceedings of the Seventh Annual Workshop on Cyber Security and Information Intelligence Research - CSIIRW ’11*. ACM Press, 2011. DOI: `10.1145/2179298.2179333`. URL: `https://doi.org/10.1145/2179298.2179333`.

[19] Charlie Jacomme and Steve Kremer. “An Extensive Formal Analysis of Multi-factor Authentication Protocols”. In: *ACM Transactions on Privacy and Security* 24.2 (Feb. 2021), pp. 1–34. DOI: `10.1145/3440712`. URL: `https://doi.org/10.1145/3440712`.

[20] Liling Cao and Wancheng Ge. “Analysis and improvement of a multi-factor biometric authentication scheme”. In: *Security and Communication Networks* 8.4 (May 2014), pp. 617–625. DOI: `10.1002/sec.1010`. URL: `https://doi.org/10.1002/sec.1010`.

[21] *2018 reform of EU data protection rules*. European Commission. May 25, 2018. URL: https://ec.europa.eu/commission/sites/beta-political/files/data-protection-factsheet-changes_en.pdf (Accessed: 17.6.2019).

[22] Kruti Sharma and Kavita R Singh. "Online data back-up and disaster recovery techniques in cloud computing: A review". In: *International Journal of Engineering and Innovative Technology (IJEIT)* 2.5 (2012), pp. 249–254.

[23] *PRINCIPLES OF CHAOS ENGINEERING*. URL: https://principlesofchaos.org/ (Accessed: 12.4.2022).

[24] Primavera De Filippi and Smari McCarthy. "Cloud computing: Centralization and data sovereignty". In: *European Journal of Law and Technology* 3.2 (2012).

[25] *Parler drops Amazon web-hosting legal challenge*. URL: https://www.bbc.com/news/technology-56270138 (Accessed: 11.4.2022).

[26] *Amazon: NOT OK - why we had to change Elastic licensing*. URL: https://www.elastic.co/blog/why-license-change-aws (Accessed: 8.4.2022).

[27] *U.S.-Russian Relations*. URL: https://www.csis.org/programs/russia-and-eurasia-program/archives/us-russian-relations (Accessed: 20.4.2022).

[28] *Russian Data Protection Laws: Essential Guide on Compliance Requirements in Russia*. URL: https://incountry.com/blog/russian-data-protection-laws-essential-guide-on-compliance-requirements-in-russia/ (Accessed: 28.4.2022).

[29] Darko Golec, Ivan Strugar, and Drago Belak. "The Benefits of Enterprise Data Warehouse Implementation in Cloud vs. On-premises". In: *ENTRENOVA - ENTerprise REsearch InNOVAtion* (2021).

[30] A Srinivas, Y Seetha Ramayya, and B Venkatesh. "A study on cloud computing disaster recovery". In: *International Journal of Innovative Research in Computer and Communication Engineering* 1.6 (2013), pp. 1380–1389.

[31] *GCP Cloud sustainability*. URL: https://cloud.google.com/sustainability (Accessed: 18.4.2022).

[32] *GCP Cloud locations*. URL: https://cloud.google.com/about/locations (Accessed: 20.4.2022).

[33] *GCp Pricing*. URL: https://cloud.netapp.com/blog/gcp-cvo-blg-google-cloud-pricing-the-complete-guide (Accessed: 20.4.2022).

[34] Parul Kudtarkar et al. “Cost-effective cloud computing: a case study using the comparative genomics tool, roundup”. In: *Evolutionary Bioinformatics* 6 (2010), EBO–S6259.

[35] Justice Opara-Martins, Reza Sahandi, and Feng Tian. “Critical analysis of vendor lock-in and its impact on cloud computing migration: a business perspective”. In: *Journal of Cloud Computing* 5.1 (Apr. 2016). DOI: 10.1186/s13677-016-0054-z. URL: https://doi.org/10.1186/s13677-016-0054-z.

[36] *Russia blocks tens of thousands of IP addresses owned by the cloud infrastructure provider DigitalOcean*. URL: https://meduza.io/en/news/2018/04/18/russia-blocks-tens-of-thousands-of-ip-addresses-owned-by-the-cloud-infrastructure-provider-digitalocean (Accessed: 28.4.2022).

[37] *Global Clouds and Cloud Providers in Russia*. URL: https://incountry.com/blog/global-clouds-and-cloud-providers-in-russia/ (Accessed: 28.4.2022).

[38] *What is IAC?* URL: https://docs.microsoft.com/en-us/devops/deliver/what-is-infrastructure-as-code (Accessed: 12.4.2022).

[39] Kief Morris. *Infrastructure as code: managing servers in the cloud.* " O’Reilly Media, Inc.", 2016.

[40] *Top 10 Infrastructure as Code (IaC) Tools to Know in 2022*. URL: https://spectralops.io/blog/top-10-infrastructure-as-code-iac-tools-to-know-in-2022/ (Accessed: 2.4.2022).

[41] *Ansible*. URL: https://www.ansible.com/ (Accessed: 12.4.2022).

[42] *Yandex Object Storage*. URL: https://cloud.yandex.com/en/services/storage (Accessed: 28.4.2022).

[43] *System Requirements*. URL: https://wiki.openmrs.org/display/docs/System+Requirements (Accessed: 16.4.2022).

[44] *Customizing OpenMRS*. URL: https://github.com/openmrs/openmrs-book-guide/blob/master/en/Configuration/customizing-openmrs-with-plug-in-modules.md (Accessed: 16.4.2022).

[45] *OpenMRS Around the World*. URL: https://archive.flossmanuals.net/openmrs-guide/ (Accessed: 16.4.2022).

[46] *OpenMRS Google Summer of Code*. URL: https://openmrs.org/category/google-summer-of-code/ (Accessed: 16.4.2022).

[47] *Google Summer of Code*. URL: https://summerofcode.withgoogle.com/about (Accessed: 16.4.2022).

[48] *OpenMRS from Scratch*. URL: https://wiki.openmrs.org/display/docs/OpenMRS+from+Scratch (Accessed: 16.4.2022).

[49] Haleh Ayatollahi, Peter A. Bath, and Steve Goodacre. "Paper-based versus computer-based records in the emergency department: Staff preferences, expectations, and concerns". In: *Health Informatics Journal* 15.3 (Aug. 2009), pp. 199–211. DOI: 10.1177/1460458209337433. URL: https://doi.org/10.1177/1460458209337433.

Appendices

Appendix 1 - Non-exclusive licence

I German Ivanov

1. Grant Tallinn University of Technology free licence (non-exclusive licence) for my thesis "A Secure Infrastructure Solution for a Plastic Surgery Clinic", supervised by Kaido Kikkas
   (a) to be reproduced for the purposes of preservation and electronic publication of the graduation thesis, incl. to be entered in the digital collection of the library of Tallinn University of Technology until expiry of the term of copyright;
   (b) to be published via the web of Tallinn University of Technology, incl. to be entered in the digital collection of the library of Tallinn University of Technology until expiry of the term of copyright.
2. I am aware that the author also retains the rights specified in clause 1 of the non-exclusive licence.
3. I confirm that granting the non-exclusive licence does not infringe other persons' intellectual property rights, the rights arising from the Personal Data Protection Act or rights arising from other legislation