THE REPUBLIC OF TURKEY
BAHCESEHIR UNIVERSITY

TURKISH PERSONAL DATA PROTECTION SYSTEM

Master's Thesis

HALIL URAL

İSTANBUL, 2020

THE REPUBLIC OF TURKEY
BAHCESEHIR UNIVERSITY

INSTITUTE OF SCIENCE AND TECHNOLOGY
COMPUTER ENGINEERING DEPARTMENT

TURKISH PERSONAL DATA PROTECTION SYSTEM

Master's Thesis

HALIL URAL

Thesis Supervisor: ASSIST. PROF. DR. ÖVGÜ ÖZTÜRK ERGÜN

İSTANBUL, 2020

T.C. BAHCESEHIR UNIVERSITY

INSTITUTE OF SCIENCE AND TECHNOLOGY

COMPUTER ENGINEERING MASTER'S PROGRAM

Title Of the thesis: Turkish Personal Data Protection System

Name and Surname of The Student: Halil URAL

Thesis Defense Day: 04/26/2020

The thesis has been approved by the Graduate School of Natural and Applied Sciences which has fulfilled the necessary conditions as Master Thesis.

Assoc. Dr. Burak KÜNTAY

Institute Director

I certify that this thesis meets all the requirements as Master thesis.

This thesis was read by us, quality and content as a Master's thesis has been seen and accepted as sufficient.

	Title/Name	Signature
Thesis Advisor's	Assist. Prof. Dr. Övgü ÖZTÜRK ERGÜN	
Member's	Assist. Prof. Dr. Tarkan AYDIN	
Member's	Assist. Prof. Dr. Gökhan İNCE	

ABSTRACT

TURKISH PERSONAL DATA PROTECTION SYSTEM

Halil URAL

Computer Engineering Department

Thesis Supervisor: Assist. Prof. Dr. Övgü ÖZTÜRK ERGÜN

June 2020, 101 pages

Personal Data Protection Law No. 6698 was enacted on March 24, 2016 in Turkey. All natural or legal persons holding personal data are covered by this law. With the enforcement of the Personal Data Protection Law, audits have started and serious steps have been taken regarding the protection of personal data. With the enactment of the law, companies must inform the person to whom they will process personal data and obtain open consent.

Companies are required to comply with the legal provisions regarding the Personal Data Protection Law. With the increase of the processed data, personal data has become a difficult problem to manage today. Many companies manage the KVKK compliance process manually with programs such as excel, word or physical documents. Companies are required to report to auditors that their KVKK processes are operating in accordance with the law, and to make quick returns to their customers asking which personal data is stored or requesting destruction.

In this thesis, concepts such as when the Personal Data Protection Law was published, who the law covers, what data is included in the scope of personal data, details of the processes of storing, transmitting and destroying personal data, the responsibilities of the companies that store personal data, and the rights of the people who use personal data will be discussed. The system will be designed to automate the manual management of companies' compliance with the Personal Data Protection Law. Thanks to this system, companies 'work force will decrease, productivity will increase and companies' KVKK process risks will be minimized.

Keywords: Personal Data, Personal Data Protection Law, KVKK

ÖZET

TÜRKİYE'DE KİŞİSEL VERİLERİ KORUMA PROBLEMİNİN
ÇÖZÜMÜ İLE İLE İLGİLİ SİSTEM YAKLAŞIMI

Halil URAL

Bilgisayar Mühendisliği Departmanı

Tez Danışmanı: Dr. Öğr. Üyesi Övgü ÖZTÜRK ERGÜN

Haziran 2020, 101 sayfa

6698 sayılı Kişisel Verileri Koruma Kanunu, Türkiye'de 24 Mart 2016 yılında kanunlanmıştır. Kişisel veri bulunduran tüm gerçek ya da tüzel kişiler bu kanun kapsamına girmektedir. Kişisel Verileri Koruma Kanunu'nun yürürlüğe girmesi ile denetimler başlamış ve kişisel verilerin korunması ile ilgili ciddi adımlar atılmıştır. Kanunun yürürlüğe girmesi ile şirketlerin kişisel veri işleyeceği kişiyi bilgilendirmesi ve açık rıza alması gerekmektedir.

Şirketlerin Kişisel Verileri Koruma Kanunu ile ilgili kanuni hükümleri yerine getirmesi gereklidir. İşlenen verinin artması ile kişisel verilerin günümüzde yönetilmesi güç bir problem haline gelmiştir. KVKK uyum sürecini birçok şirket excel, word gibi programlarla ya da fiziksel dokümanlarla manuel olarak yönetmektedir. Şirketlerin denetçilere KVKK süreçlerinin kanuna göre işlediğini raporlamaları ve hangi kişisel verilerinin saklandığını soran ya da imha edilmesini isteyen müşterilerine hızlı dönüş yapması gerekmektedir.

Bu tezde, Kişisel Verileri Koruma Kanunu'nun ne zaman yayınlandığı, kanunun kimleri kapsadığı, kişisel veri kapsamına hangi verilerin girdiği, kişisel verilerin saklama, iletme ve imha süreçlerinin detaylandırılması, kişisel veri saklayan şirketlerin sorumlulukları, kişisel verisi kullanılan kişilerin hakkları gibi kavramlar ele alınacak ve şirketlerin Kişisel Verileri Koruma Kanunu'na uyum süreçlerinin manuel yönetimini otomatize etmek için sistem tasarlanacaktır. Bu sistem sayesinde şirketlerin iş gücü azalacak, verimlilik artacak ve şirketlerin KVKK süreç riskleri minimuma indirgenecektir.

Anahtar Kelimeler: Kişisel Veri, Kişisel Verilerin Koruma Kanunu, Kvkk

CONTENTS

TABLES

FIGURES

ABBREVIATIONS

DAAS	:	Database as a Service
GDPR	:	General Data Protection Regulation
JSON	:	JavaScript Object Notation
KVKK	:	Kişisel Verileri Koruma Kanunu
MTD	:	Multi-Tenant Database
SFTP	:	Secure File Transfer Protocol
SQL	:	Structured Query Language

1. INTRODUCTION

Law No. 6698 on Personal Data Protection was published in the Official Gazette on 24 March 2016 and entered into force. The purpose of the law is to protect the fundamental rights and freedoms of individuals, especially the privacy of personal life, in the processing of personal data, and to regulate the obligations and procedures and principles to be followed by real and legal persons who process personal data (Kişisel Verileri Koruma Kanunu, 2016).

The processing of personal data is any kind of information about a specific or identifiable natural person (Kişisel Verileri Koruma Kanunu-Madde3.d), and the processing of personal data is obtained, stored, changed and re-created by the personal data, which is fully or partially automated or non-automated provided that it is part of any data recording system. It refers to any action taken on data such as regulation, disclosure, transfer, takeover, making it available, classifying or preventing its use (Kişisel Verileri Koruma Kanunu, 2016).

The law defines the concept of personal data processing as “real or legal person who processes personal data on behalf of him/her based on the authority given by the data controller”. With the enforcement of the law, those who process personal data must inform the personal data owner and obtain open consent and companies needed to make changes in the processes of storing, transmitting and destroying personal data.

Almost every system used by companies today has personal data. With the advances in technology, data sharing has increased and personal data has become a difficult problem to manage. Many companies manage the KVKK compliance process manually with programs such as excel, word or physical documents and they are required to report to auditors that their KVKK processes are operating in accordance with the law, and to make quick returns to their customers asking which personal data is stored or requesting destruction.

In this thesis, concepts such as when the Personal Data Protection Law was published, who the law covers, what data is included in the scope of personal data, details of the processes of storing, transmitting and destroying personal data, the responsibilities of

the companies that store personal data, and the rights of the people who use personal data will be discussed. The system will be designed to automate the manual management of companies' compliance with the Personal Data Protection Law. Thanks to this system, companies 'work force will decrease, productivity will increase and companies' KVKK process risks will be minimized.

2. LITERATURE REVIEW

2.1 PERSONAL DATA PROTECTION LAW

With the rapidly developing technology, the collection, sharing, processing and storage of the data have increased. The confidentiality of private life, which is one of the fundamental rights and freedoms, is directly related to the protection of personal data. The developing technology has brought security risks related to data protection. The data that companies may have inconvenience in the possession of third parties should be protected legally.

Today, both public and private institutions can access various information about thousands of people every day. The information obtained can be easily processed and transferred due to the developments in information technologies. As a result, there was a need to protect personal data in (KVKK, 2018).

As in the rest of the world, measures to protect personal data have been taken in our country and Personal Data Protection Law No. 6698 was published in the official newspaper on March 24, 2016. The main factors that lead our country to prepare a legal regulation for the protection of personal data; Effective protection of human rights, membership negotiations with the European Union and international cooperation and the need to increase trade (KVKK 2018).

2.1.1 International Standards In Protection Of Personal Data

With the acceptance of the protection of personal data as a legal right, many international regulations and studies have been made.

For the first time in 1948, in the Universal Declaration of Human Rights for the privacy and protection of private life, “No one can be exposed to arbitrary confusion, rape against honor and fame regarding his private life, family, residence or correspondence. Everyone has the right to be protected by law against these conflicts and rape (Universal Declaration of Human Rights, 1948).

With the provisions of Article 8 of the European Convention on Human Rights, which came into force in 1953, the right of individuals to privacy and privacy is legally guaranteed.

Article 8 of the European Convention on Human Rights - Right to respect for private life and family life (European Convention on Human Rights 1950).

a. Everyone has the right to respect for his private life and family life, residence and communication.

b. The exercise of this right is made by a public authority, for the benefit of national security, public security or the economic well-being of the country, for the prevention of crime or disorder, for the protection of health or morality, or for the protection of the rights and freedoms of others, and a democratic There will be no intervention, except those deemed necessary in the community.

The first national data protection law is the Hessen State Data Protection Law of Germany in 1970. This Law has been prepared with the aim of determining the procedures and principles regarding the acquisition and storage of data in the face of access to land records with the help of information systems (KVKK 2018).

The two resolutions adopted by the Council of Europe in 1973 and 1974 as an international regulation to set the necessary standards for the protection of personal data held in private and public electronic databases resulted in subsequent regulations on the protection of personal data (KVKK 2018).

The first comprehensive international agreement on the protection of personal data was the "Convention on the Protection of Persons Against Automatic Processing of Personal Data" dated 1981 and numbered 108 of the Council of Europe. The Committee of Ministers of the Council of Europe has issued a total of 13 recommendations, stating the procedures and principles for the implementation of Convention 108 (KVKK 2018).

Following these developments, while national legislation was established in the European countries and the USA, various directives and international agreements were prepared within the scope of the United Nations, Council of Europe, Organization for

Economic Cooperation and Development (OECD) and the European Union (KVKK 2018).

Finally, GDPR - General Data Protection Regulation - was adopted by the European Union on 25 May 2018. GDPR covers not only EU member states but also all countries that exchange personal data with a person residing in EU member states (ÜNAL NİZAMOĞLU, 2018).

2.1.2 Terms Of Protection Of Personal Data Law

2.1.2.1 Persons

2.1.2.1.1 Data Officer

The real or legal person who carries out his/her personal data processing activity is defined as the data officer. Data Officers;

a. Determining the purposes and means of processing personal data,

b. Setting up and managing a recording system,

c. It is held responsible for taking all kinds of technical and administrative measures regarding the protection of personal data.

2.1.2.1.2 Contact Person

If the data officer is a legal entity, T.C., who will contact the Personal Data Protection Authority and the Personal Data Owners. the natural person who is a citizen is also defined as a liaison person who should be the contact person should be stated to VERBİS during registration.

2.1.2.1.3 Data Subject

The data subject is the real person whose Personal data is processed.

2.1.2.1.4 Data Processor

Based on the authority given by the data officer, the real or legal person who processes personal data on his/her behalf is called “Data Processor”.

2.1.2.1.5 Personal Data

All kinds of information about an identified or identifiable natural person is defined as “Personal Data”. Identity, Communication, Location, Personality, Legal Transaction, Customer Transaction, Physical Space Security, Finance, Visual and Audio Records, Marketing, Professional Experience data are personal data.

2.1.2.1.6 Special Qualified Personal Data

If it is learned by its qualifications, personal data that can cause the person concerned to be victimized or subject to discrimination is defined as Special Qualified Personal Data.

Race, Ethnicity, Political thought, Religion, Sect or other beliefs, Disguise and dress, Data on association, foundation or union membership, Health, Sexual life, Criminal conviction and security measures, Biometric and genetic data are special personal data.

2.1.2.1.7 Express Consent

It expresses the consent of a particular subject, which is based on information and explained with free will.

2.2 PERSONAL DATA PROTECTION LAW PROCESSES

2.2.1 Anonymize

It defines the process of making personal data unrelated to an identifiable or identifiable natural person under any circumstances, even by matching with other data.

2.2.2 Personal Data Record System

The acquisition, recording, storage, preservation, modification, reorganization, disclosure, transfer, acquisition, availability, classification, or use of personal data, whether fully or partially automated, or non-automated, provided that it is part of any data recording system. Any transaction performed on data such as blocking means processing of personal data.

2.2.3 Creating Personal Data Inventory

One of the most important stages of the process of compliance with Personal Data Protection law is the determination of the existing "Personal data" inventory. To create the data inventory, you need to determine which personal data you receive and store for which activities in your workplace (eg human resources recruitment, sales customer information, staff personal processes, etc.). Create your personal data inventory by matching your work and contact groups.

2.2.4 Preparing Kvkk Discipline Regulations For Personnel

Personnel have important duties in compliance with personal data protection law. If the personnel do not comply with the security policies and procedures, you should create the discipline regulations and share them with your personnel. When you save the discipline regulation and make its status published, it will be announced to all your staff.

2.2.5 Reducing Personal Data As Possible

According to the Law on Protection of Personal Data, it is said that "personal data should be accurate and up-to-date, if necessary, and maintained for the period required by the relevant legislation or for the purpose for which they are processed". If you are an institution or company that has been operating for a long time, some of the personal data may become inaccurate, outdated and do not serve any purpose since you may have

collected a lot of personal data. For processing purposes, it is necessary to evaluate whether the said personal data is still needed and to ensure that the personal data is kept in the correct place. In addition, although it is suitable for personal data processing purposes in order to prevent unauthorized access, personal data, which is not frequently required and kept for archive purposes, is kept in a safer environment, and personal data is not deleted and destroyed with personal data retention and destruction policy. or be anonymized and safely destroyed.

2.2.6 Setting A Personal Data Protection Policy

You should establish "accurate and consistent" policies and procedures regarding personal data security and integrate them in accordance with your work and functioning. The level of personal data security is not sufficiently ensured when policies and procedures cannot be prepared well and on time, problem areas cannot be identified or existing security measures cannot be used. Good incident management, in which the measures to be taken in this context are predetermined, will also reduce the pressure on your staff. Therefore, you need to make sure which personal data is contained in the data logging systems and by examining the current security measures, it complies with other legal obligations. Within the scope of policies and procedures; checks should be done regularly, the checks should be documented, the issues to be developed should be determined, and checks should be continued regularly after the necessary updates are carried out. In addition, the risks that may arise for each personal data category and how security violations will be managed should be clearly determined.

2.2.7 Creating Personal Data Retention And Destruction Policy

According to the law numbered 6698, personal data must be stored at the times specified in the relevant legislation or in accordance with the processing purposes. When establishing the retention and disposal policy, you must identify the legal reasons for processing, the processing purposes and the reasons that require destruction for each personal data group. You should also state the technical and administrative measures

taken for retention within this policy. You must determine the storage and destruction times.

2.2.8 Creating Privacy Agreement

You should create confidentiality agreements regarding personal data security with the contact groups and share them with the contact groups. When you save the confidentiality agreement and make its status published, it is automatically shared with your staff.

2.2.9 Provide Training On Personnel Data Protection Law And Data Security

In case of attacks and cyber security threats that would damage personal data security, it is important for the personnel to make the first intervention, even if they have limited information, to ensure personal data security. However, issues such as unlawful disclosure or sharing of personal data are major personal data security violations. These violations can also arise from the use of weaknesses of users, such as carelessness, thoughtfulness, or inexperience, by opening an e-mail attachment containing malware or opening personal mail to third parties by sending the e-mail to the wrong recipient. For this reason, it is very important for the employees to receive training on issues such as not disclosing and sharing the personal data illegally, and to raise awareness for the employees in order to ensure data security. In case of significant changes in policies and procedures regarding personal data security; With these new trainings, it should be ensured that these changes are presented to the information of the employees and the information about the threats related to personal data security is kept up to date.

2.2.10 Review All Contracts Containing Personal Data

You should review all contracts with personal data available at your institution and ensure that they comply with the 6698 numbered personal data protection law. The contents of the contract should be special for institutions and companies according to

their main activity areas. For this reason, it is much more necessary to make arrangements with a legal advisor.

it will be healthy.

2.2.11 Preparing Lighting Texts For Relevant Persons

Law No. 6698 imposes a "Lighting Obligation" for Data supervisors who process the data of the whole person concerned. This is also a right for natural persons whose Personal data is processed. The Lighting Obligation refers to information about the personal data processed. The obligation to enlighten is not an obligation at the request of the person concerned, and the obligation to demonstrate belongs to the data controller. In case of failure to fulfill, there is an administrative fine. The obligation to enlighten One of the important steps for compliance with personal data protection law and must be followed. You must prepare different lighting texts for each contact person whose data is processed. You can easily prepare the lighting texts by using their menus and by means of visitors, exhibitors, lighting panels you can publish.

2.2.12 Setting Up A System For Open Consent

In some cases, you may need to obtain the "Open Consent Statement" of the person you are processing the Personal data for. It is determined in the law in which cases you should get an Open consent declaration. If you are processing special quality personal data and you are not one of the institutions specified in the law for this process, you should get a clear consent statement. Similarly, domestic or international data, The transfer is also possible with the explicit consent of the person concerned.

2.2.13 Identifying Information Security Risk Measures

It takes place in information servers, personal computers, cloud, portable media, files, folders, your employees' memories and in many different environments. Risk management is extremely important to ensure the security of these environments, which we call information assets. Determining what information assets you have allows you to

detect threats related to these assets in advance, both to measure risks and to take measures of high-grade risks. Every precaution you take is at risk minimizes the damage you can see.

2.2.14 Creating An Internal Audit Procedure For The Protection Of Personal Data

As a data controller, you should make or have inspections on systems containing personal data. In these audits, you should audit each department that processes personal data separately, as well as the IT department for technical measures, and the management representative for processes, policies and procedures. Audit procedure how to conduct audits, frequency of audits and who is responsible for audits should be.

2.2.15 Planning And Reporting Audits

Audits should be carried out within a plan, and auditors should be prepared for what they should audit during the audit. For this, audit plans and question lists should be prepared. It is checked together with the evidence whether the necessary measures are taken during the audits and whether the personal data are protected in accordance with the storage conditions and rules. Audit results should be reported, deficiencies with corrective actions should be followed.

2.2.16 Data Breach Information Procedure Identification

In the event that personal data processed according to the personal data protection law are obtained by others in an illegal way, the data officer must notify this to the Board within 72 hours at the latest without delay. In addition, determining the people affected by the data violation in question, if it can directly contact the relevant people's contact addresses, if not, by sharing this situation on their website and announcing it to the relevant people. It is legal obligation.

2.2.17 Take Technical Measures To Protect Your Information Technology Systems

The view that all of these measures can be taken using a single security product is not always true. Cyber security measures need to be considered as a whole, from unauthorized access threats from the internet, backing up, password policies, access rights to environments containing personal data, e-mail security, and monitoring of transactions.

2.2.18 Authority Matrix

Access to systems containing personal data should be limited. In this context, employees should be granted access to the extent necessary for their work and duties, as well as their powers and responsibilities, and an access authorization and control matrix should be created. Authority matrices which users / groups have access to which files, what are the write-read-record deletions, which users / groups it should contain information such as using software.

2.2.19 Recording Logs (Traces) Of Accessing Personal Data

Recording all transaction transactions of all users is extremely important for the possibility of encountering any data breach. Access logs should contain information about user, transaction time, IP information, personal data accessed, if there is any change, before and after the change, and reporting of this data. Access logs also include software logs, 5651 hotspot accesses and device logs. Bilimp offers a wide log structure in itself.

2.2.20 Creating A Firewall

Firewall devices are used to prevent attacks from outside, to prevent unauthorized people from sending information to outside, to reduce virus threats and to limit unauthorized internet access. The firewall is correctly configured and the software in it is up to date

Keeping it is at least as important as having this device.

2.2.21 Backing Up Data

During the storage period required by your processing purpose, you should keep your personal data and set up a backup system that will allow you to access this data when needed. It is not enough to take backups to your servers within the company, it is not enough to periodically backup these backups to a different environment.

2.2.22 Data Information Sets

Personal data has sets and subsets, it is important for grouping personal data. Below is a list of data information sets and a description in the column next to it. Especially when creating data inventory, it is necessary to use data information sets and sub-sets that depend on them, so that which data inventories. It can easily be found in the system, which is associated with personal data.

Table 2.1: Example Data Informatin Sets in KVKK

Doğum günü	Doğum günü
Kimlik	Bu veri kategorisi Ad soyad, Anne - baba adı, Anne kızlık soyadı, Doğum tarihi, Doğum yeri, Medeni hali, Nüfus cüzdanı seri sıra no, TC kimlik no gibi veri türlerini ifade etmektedir.
İletişim	Bu veri kategorisi Adres no, E-posta adresi, İletişim adresi, Kayıtlı elektronik posta adresi (KEP), Telefon no gibi veri türlerini ifade etmektedir.
Lokasyon	Bu veri kategorisi Bulunduğu yerin konum bilgileri gibi veri türlerini ifade etmektedir.
Özlük	Bu veri kategorisi Bordro bilgileri, Disiplin soruşturması, İşe giriş belgesi kayıtları, Mal bildirimi bilgileri, Özgeçmiş bilgileri, Performans değerlendirme raporları gibi veri türlerini ifade etmektedir.
Hukuki İşlem	Bu veri kategorisi Adli makamlarla yazışmalardaki bilgiler, Dava dosyasındaki bilgiler gibi veri türlerini ifade etmektedir.
Müşteri İşlem	Bu veri kategorisi Çağrı merkezi kayıtları, Fatura, senet, çek bilgileri, Gişe dekontlarındaki bilgiler, Sipariş bilgisi,

	Talep bilgisi gibi veri türlerini ifade etmektedir.
Fiziksel Mekan Güvenliği	Bu veri kategorisi Çalışan ve ziyaretçilerin giriş çıkış kayıt bilgileri, Kamera kayıtları gibi veri türlerini ifade etmektedir.
İşlem Güvenliği	Bu veri kategorisi IP adresi bilgileri, İnternet sitesi giriş çıkış bilgileri, Şifre ve parola bilgilerigibi veri türlerini ifade etmektedir.
Risk Yönetimi	Bu veri kategorisi Ticari, teknik, idari risklerin yönetilmesi için işlenen bilgiler gibi veri türlerini ifade etmektedir.
Finans	Bu veri kategorisi Bilanço bilgileri, Finansal performans bilgileri, Kredi ve risk bilgileri, Malvarlığı bilgileri gibi veri türlerini ifade etmektedir.
Mesleki Deneyim	Bu veri kategorisi Diploma bilgileri, Gidilen kurslar, Meslek içi eğitim bilgileri, Sertifikalar, Transkript bilgileri gibi veri türlerini ifade etmektedir.
Pazarlama	Bu veri kategorisi Alışveriş geçmişi bilgileri, Anket, Çerez kayıtları, Kampanya çalışmasıyla elde edilen bilgiler gibi veri türlerini ifade etmektedir.
Görsel Ve İşitsel Kayıtlar	Bu veri kategorisi Görsel ve İşitsel kayıtlar gibi veri türlerini ifade etmektedir.
Irk Ve Etnik Köken	Bu veri kategorisi Irk ve etnik kökeni bilgileri gibi veri türlerini ifade etmektedir.
Siyasi Düşünce Bilgileri	Bu veri kategorisi Siyasi düşüncesini belirten bilgiler, Siyasi parti üyeliği bilgisi gibi veri türlerini ifade etmektedir.
Felsefi İnanç, Din, Mezhep Ve Diğer İnançlar	Bu veri kategorisi Diğer inançlarına ilişkin bilgiler, Dini aidiyetine ilişkin bilgiler, Felsefi inancına ilişkin bilgiler, Mezhep aidiyetine ilişkin bilgiler gibi veri türlerini ifade etmektedir.
Kılık Ve Kıyafet	Bu veri kategorisi Kılık kıyafete ilişkin bilgiler gibi veri türlerini ifade etmektedir.
Dernek Üyeliği	Bu veri kategorisi Dernek üyeliği bilgileri gibi veri türlerini ifade etmektedir.

2.3 PREVIOUS STUDIES ON PERSONAL DATA AND PRIVACY PROTECTION:

Since personal data protection law is of great importance for companies and it is difficult to manage manually, systematic perspectives are important for companies. Because, in cases where the kvkk processes cannot be managed properly, there is a strong and complete kvkk because of the legal transactions and fines and management application should acquire (Dawoud, 2006).

There are also studies on this subject in cloud systems. Recently, cloud computing systems are seen as one of the most important developments in the field of Information Technology (IT) applications. The data stored and processed in the cloud system is increasing rapidly with each passing day. Its usage is becoming widespread, as the data is accessible from every point, increases efficiency thanks to central data storage and central processing capacity, saves many applications from high infrastructure costs, and offers flexible and scalable solutions. In the thesis work, many techniques known as PPDR (Privacy Preserving Data Retrieval) are already proposed. 9 security points have been defined in order to ensure high level of security and privacy. The technique proposed to perform the security and privacy of the data call by meeting the 9 + 1 criterion was made with a multi-server structure produced from a Radio Frequency Identification (RFID) privacy system called HEADA (Dawoud, 2006).

Another approach to protect personal data deprivation is to protect individuals' genome sequencing data. The safety of genomes, which can be shared on public websites, is present as risk. In the study, firstly, current studies on genomic secrecy attack are being developed by using the recombination model between haplotypes, kinship relationships and phenotypic features, and in addition, while protecting the privacy of individuals, genomic There is a different privacy based software framework for sharing data (Mobayenjarihani, 2018).

2.4 PRIVACY PRESERVING DATA MANAGEMENT

Database security has generally focused on access control: Users are explicitly (or perhaps implicitly) allowed certain types of access to a data item. This includes work in multilevel secure database as well as statistical queries (Adam, Wortmann 1989).

Privacy is a more complex concept. Most privacy laws balance benefit vs. risk (Duncan, Keller-McNulty, Stokes 2001): access is allowed when there is adequate benefit resulting from access. An example is the European Community directive on data protection which allows processing of private data in situations where specific conditions are met. The Health Insurance Portability and Accountability Act in the U.S. specifies similar conditions for use of data. Individual organizations may define their own policies to address their customers' needs. The problems are exacerbated in a federated environment. The task of data integration itself poses risks, as revealing even the presence of data items at a site may violate privacy.

Some of the privacy issues have been addressed for the case of a single database management system in Hippocratic Databases (Agrawal et al, 2002). Other privacy issues have been addressed for the case of a single interaction between a user and a Website in the P3P standard (Cranor et al, 2002). None of the current techniques address privacy concerns when data is exchanged between multiple organizations, and transformed and integrated with other data sources.

There are a lot of studies on privacy management. As an example, consider secure supply-chain management (Atallah et al, 2003). An example scenario would be two companies that use a common raw material. Knowing that they share this need and coordinating their orders and production would enable smoothing out the supply line and improving overall supply chain efficiency. A prerequisite for this coordination is the ability to identify the common raw material, suppliers, customers, etc., without giving up competitive knowledge advantages or violating anti-trust law. Standards for sharing logistics information cover such a wide ground that ambiguity is inevitable (e.g., the ECCMA Open Technical Dictionary has over 30,000 standard attribute names).

Data Integration and Data Mining are quite closely coupled. Integration is a necessary pre-requisite before mining data collected from multiple sources. At the same time, data mining/machine learning techniques are used to enable automatic data integration. Several systems have been developed to implement automatic schema matching (Li et al, 2000). The systems use machine learning/data mining tools to help automate schema matching. SemInt (Li et al, 2000) uses neural networks to determine match candidates. Clustering is done on similar attributes of the input schema. The signatures of the cluster centers are used as training data. Matching is done by feeding attributes from the second schema into the neural network. LSD (Doan et al,2003) also uses machine learning techniques for schema matching. LSD consists of several phases. First, mappings for several sources are manually specified. Then source data is extracted (into XML) and training data is created for each base learner. Finally the base learners and the meta-learner are trained. Further steps are carried out to refine the weights learned. The base learners used are a nearest neighbor classification model as well as a Na¨ıve Bayes learner. Again, there has been work on different privacy-preserving classification models (Vaidya et al, 2004) that is applicable. Artemis (Castano et al, 1999) is another schema integration tool that computes “affinities” in the range 0 to 1 between attributes. Schema integration is done by clustering attributes based on those affinities. Clearly, a lot of work in both privacypreserving data mining as well as cryptography is relevant to the problem of privacy-preserving schema integration. However, it is not yet clear how this could be applied efficiently. Record linkage also uses various machine learning techniques. Record linkage can be viewed as a pattern classification problem (Elfeky et al, 2002). In pattern classification problems, the goal is to correctly assign patterns to one of a finite number of classes. Similarly, the goal of the record linkage problem is to determine the matching status of a pair of records brought together for comparison. Machine learning methods, such as decision tree induction, neural networks, instance- based learning, clustering, are widely used for pattern classification. Given a set of patterns, a machine learning method builds a decision model that can be used to predict the class of each unclassified pattern. Again, prior privacy-preserving work is relevant. At the other end of the spectrum, privacypreserving data mining assumes that data integration has already been done, which is clearly not a solved problem.

Cloud computing is a model of information processing, storage, and delivery in which highly centralized physical resources are furnished to remote clients on demand. Rather than purchasing actual physical devices, servers, storage, and networking equipmen, clients lease these resources from a cloud provider as a outsourced service that abstracts away physical devices.

3. DATA AND METHOD

It is designed to manage the KVK processes, the related system is actually a corporate web application, each customer using the KVKK system will log into their own systems with a different subdomain with tenant logic, db side, sftp side are completely isolated from each other will be separated.

3.1 MULTITENANT DB APPROACH

All tenants, ie customers, should have schemas isolated from each other. In this way, all the data of the customers' kvkk processes are stored in their databases in isolation from each other with isolated schemes. This structure, called database as a service (DaaS), is actually to provide a database as a service (Moreton, Dudley, Olumuyiwa, 2014).

Walraven et al. Applied the Multi-tenancy approach at the infrastructure, operating system, middleware or even application level (Walraven, Joosen, Handekyn, Truyen, Van Landuyt 2014). Multi-tenancy db approach has become the architecture that can be used in many software, even layers such as operating system level, infrastructure level can be given as examples. In addition, isolation and resource use are among other features in the multitenancy approach. As in Figure 3.1, the Multitenancy approach has
4 main categories, these are isolated, infrastructure, common tenancy and implementation. Isolated tenancy argues that it is a tenant's own database instance additionally requires deployment. In Infrastructure tenancy, each tenant shares the same infrastructure, tenants have multiple application and database layers in terms of application and database, however, they all have to share the same hardware, so it is necessary to pay attention to system resources when working with such an

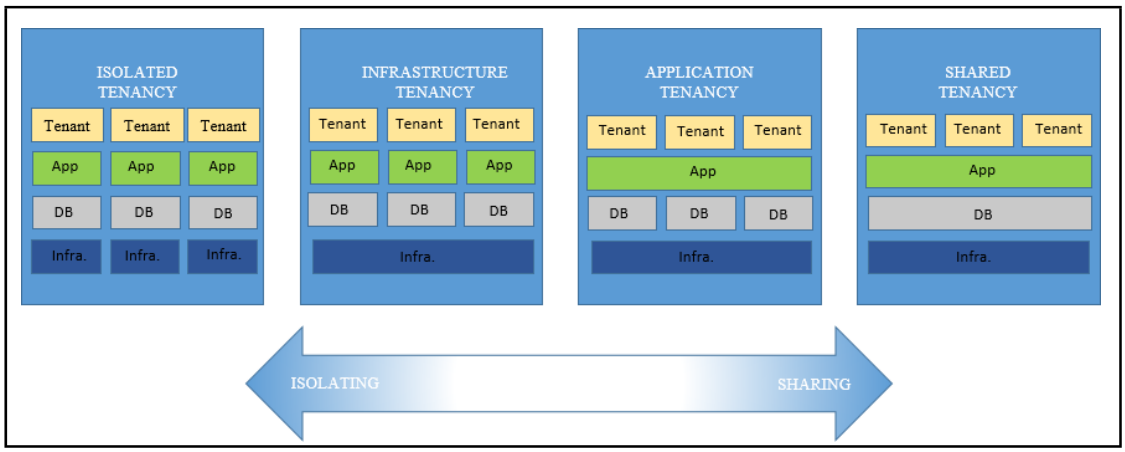
infrastructure. Application tenancy is each tenant, It is limited to the ability to have one database, they cannot have more than one database, as in infrastructure tenancy, the resources of the host on which it is located are equal and they share in a qualified way.

Finally, in shared tenancy, each tenant will have the same application and the same database, they will be added to each table as tenant_id or a different composite key within a single related database. This sharing is the easiest and first approach. As seen in Figure 3.1, going towards the right side of the picture, the tenancy approach in multi- tenancy is growing.

In addition, as you move to the right in the diagram, the sharing of resources used by tenants increases and all resources, which are described as application resources or database resources on the hardware, are shared more.

Contrary to the left side of the diagram, the isolation increases more and the sharing decreases because the application resources, database resources and infrastructure resources of each tenant may be different.

Figure 3.1: Multitenancy's big picture



As seen in Figure 3.1, the content of each approach has different approaches in terms of different usage areas. According to these approaches, if the developer of the system wants more isolation, that is, as much as possible, the common use of resources, isolated tenancy or infrastructure tenany. should adopt the approach, but if the common

use of application and database resources is required as much as possible, the approaches that will be needed here will be to choose the application and the most pure approach, shared tenancy.

In addition, if the shared tenancy approach is used, software and hardware costs will be reduced to a minimum because the application and infrastructure are the same.

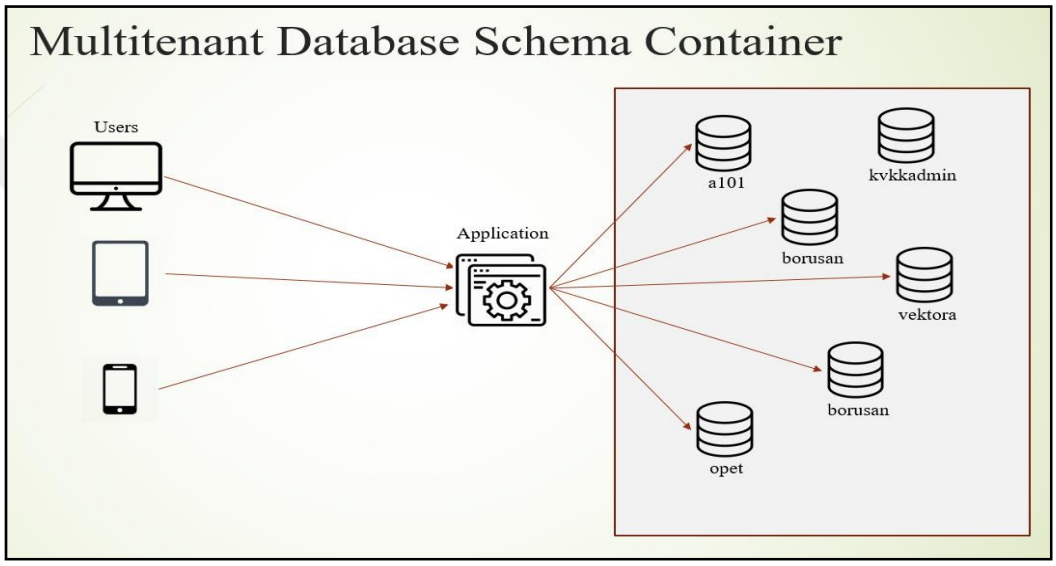
Since the database and application are single, code enhancements and database changes will be made more easily. Not to ignore the advantages of isolated tenancy must. To this extent, according to the diagram, it is much easier to add tenants in isolated tenancy because system resources are scalable as much as possible because different infrastructure and different databases are used.

has a structure. Of course, given that the code changes in the application layer, especially the build, test and deployment processes, each application will go through these processes here.

effort increases inevitably. Likewise, since the databases are separate from each other, effort will be maximized in database updates. In Shared Tenancy, there may be difficulties in data security point, since tenants are kept in the same database, composite key and share the same database (Banville, Holzel, 2012).

Predefined constants when the user is using the application are for example: / It is in the table in the data indicating the forms of appeal such as male and female. This table is called a master table and if you want to use a new customer product, it is replicated from this table. Therefore, all predefined data, except for the customer's own data, is transferred once to the customer's db.

Figure 3.2: Multitenant Database Schema Container



A db design is intended in the same way as a projection to the image above. Since the databases are completely isolated from each other, there will be no situation in a customer's database, so the other databases will not be affected because they are isolated and discrete. In addition, an automatic job is running for daily backups of databases.

Customers are authorized only to their own databases, so customer data isolation has been provided, and in case of need, specific needs are provided for each customer, and this approach provides great benefits to the system when considering the indispensable needs of databases such as Copying, Backup and Monitoring. (Replication, backing up, and monitoring can be set up on the catalog-level, hence all schemas could benefit from it. Firstly, it was aimed to run customer databases on the same hardware, but this Depending on the structure, databases can be easily adjusted to run on different hardware at any time.

3.1.1 Implementation In Multi Tenancy Approach

The multi-tenant model is very important for service providers who use Saas architecturally, because the main foundation is based on this. When designing these structures, providers develop cloud systems that include features such as accepting the tenant, allocating resources. It usually bases them on a subscription basis, that is, a new tenant to receive services can share the provider's resources (database, infrastructure), etc. after subscribing. And you can also use every service that comes with the SaaS model, which makes their job easier, and this is the KVKK application, because every tenant who wants to use the KVKK application subscribes, a schema is allocated to it, the frontend application is copied according to the subdomain, now after that is to realize the tenant's lease processes. Therefore, each tenant in the KVKK application shares different schemas on the same database, again on the same infrastructure.

According to Jacobs and Aulbach, in line with the multi-tenancy approach, there are 3 different approaches, each tenant can share the same machine, in the same application layer.

processes can be operated, and finally data can be stored on the same table (Jacobs, Aulbach 2007). As a disadvantage of these approaches, isolation between tenants decreases further, resource sharing and hardware sharing increases accordingly, the performance of the system decreases as the tenant is added, and in addition, problems such as data security arise. In response, Grund evaluated these problems and evaluated and compared each of these approaches within itself (Jacobs, Aulbach 2007), shared resources for tenants, namely server, application and database, were evaluated and their effects were investigated at the server level (Grund, 2018). As part of the processes, tenants share the same database, each tenant has one schema, the database is physically the same, but different schemes are shared within the database, and the resource is shared more efficiently in this way. However, when there is a large amount of data in one scheme, database service may be difficult because it is in the same physical database and this approach may affect other tenants. In the multi-tenancy process that is done by sharing a table, the physical database is the same, the process is carried out on only one scheme and tenants are shared as a column in the tables.

According to Schiller, when these three approaches are evaluated, on the shared server when the resources are shared by tenants. A database has been shown that only a few tenants are allowed (Schiller, 2011). In the shared resources approach, the memory spent per tenant becomes less, in other words, since each tenant has its own scheme, the main memory consumption on the server will naturally increase as the number of tenants increases.

In addition, as the number of tenants increases, the memory consumption remains steady in the structure where tenant is kept within the table. Again, according to Schiller, the shared table approach is a little longer and as such, it has a better point when we consider other approaches, as one of the reasons, it uses a minimum overhead per tenant and hence this structure is more suitable for systems with a high number of tenants, but the number of data needs to be small, up to 50MB per each tenant this is appropriate for tenants who have the number of data.

Based on all previous research, Elmore has researched, evaluated and compared all these models within the context of multi-tenancy (Elmore, 2011).

These scopes were evaluated at the point of isolation and sharing of resources, and this study was carried out to determine what is the most appropriate and efficient way when designing a system. As seen in Table 3.1, cloud evaluations were made in the context of isolation and resource sharing and the results are shown (Elmore, 2011).

Table 3.1: Cloud architectural approaches within the scope of multi-tenancy.

#	Sharing Model	Isolation	IaaS	PaaS	SaaS
1	Shared Hardware	VM	✓	✓	
2	Shared VM	OS User		✓	
3	Shared OS	DB Instance		✓	
4	Shared Instance	Database		✓	

5	Shared database	Scheme		✓	
6	Shared table	Row		✓	✓

3.1.2 Mysql Scripts Files

Use of the MySQL client can be improved at times by using script files. These files are normal text files with a .sql file extension. They can provide speed, reuse, and edibility improvements over the command line mysql client.

Script files contain any MySQL client-readable commands that could be directly invoked on the interactive client. Each statement can be separated by a line break, and terminated by semicolons (;). Script files can be used in two different ways. Within the interactive mysql client, script files can be called as any other MySQL client command would be invoked:

Figure 3.3: Multitenant Database Schema Container

```
mysql -u root -p < C:\Users\User\Desktop\filename.sql
```

Two sql scripts are used in the system, one is used for initial database operations, and the other includes predefined data that will be added there after tenant databases are created.

Figure 3.4:Initial database operation SQL(TODO)

```
mysql --host=localhost --protocol=TCP --port=3360 --user=root --password=root --
default-character-set=utf8 < "%CD%"\kvkk-vektora-db\process\initial_db_operation.sql
```

Figure 3.5: Predefined data additional SQL(TODO)

```
mysql --host=localhost --protocol=TCP --port=3360 --user=root --password=root --
default-character-set=utf8 < "%CD%"\kvkk-vektora-db\process\add_data_operation.sql
```

3.2 DOCKER

Docker is an open source virtualization platform developed for software developers and systematics. With Docker, you can run Linux and Windows virtual containers (machines) on Linux, Windows and MacOSX. Thanks to this platform, you can easily install, test and deploy web systems. Perhaps the most important feature was "It was working on my computer, why didn't it work on the server?" is eliminating the problem.

Applications are expressed as services in containers, called containers, and almost all application types are kept as images in the docker hub. Since containers are built as images, almost every image has its own minimized operating systems, and containers operate on these operating systems, so it is independent of the operating system the docker is operating as a service. This approach is as efficient as possible in terms of performance and isolation.

it has a structure (Boettiger, 2015).

Docker offers excellent features in application development, as it has changed the way the applications are run on a traditional server basis lately, as it introduces a new microservices structure and enhances image sharing among developers. It also offers the possibility of creating a micro service and controlling the relevant micro service at the level of minimized micro operating systems. In addition, application developers can

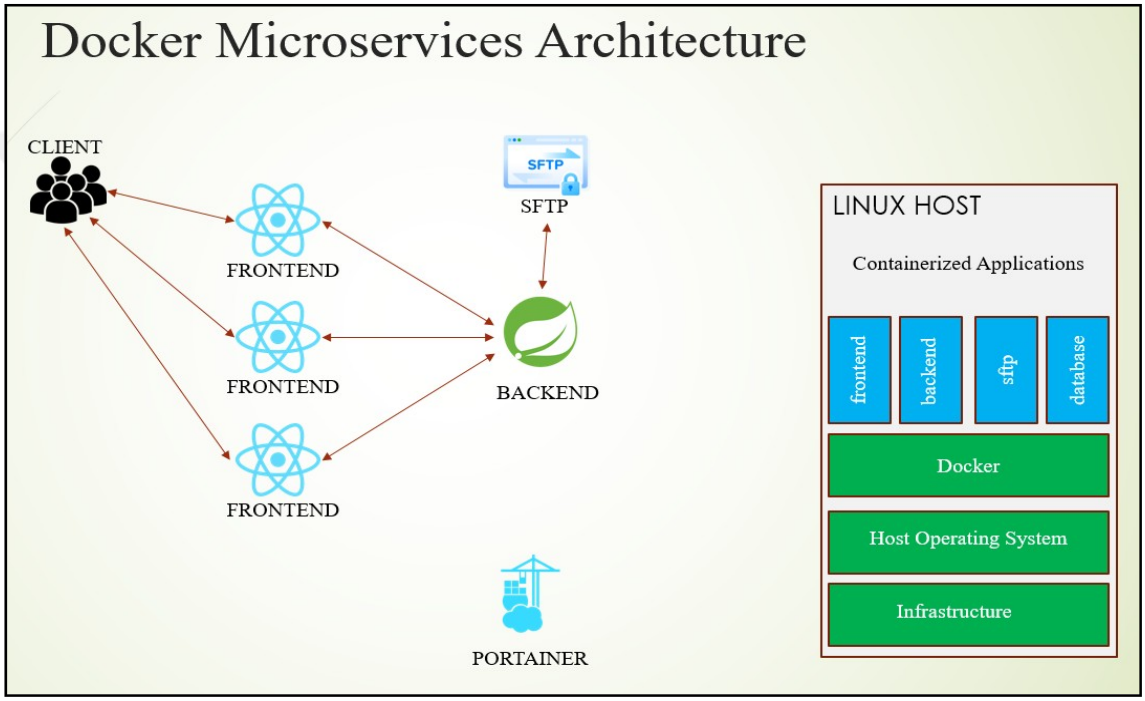
You can convert it to docker images, push it to the docker hub, and other developers can use them comfortably on their servers. Since the docker, which is a virtualized approach, runs its containers independently from the server on which it is running, the server is no longer able to work on every server, just like java applications, so micro

services running in a linux environment can easily transfer to windows servers. Today, when cloud architecture is developed, the docker is at a critical level is becoming important because it is possible to move docker services to cloud environments easily (Bui, 2015).

Docker's virtualization structure does not have a Hypervisor layer unlike known virtual machines (VirtualBox, Vmware etc.). Instead, it accesses the host operating system through Docker Engine and uses the system tools shared. Thus, it consumes less system resources than conventional VMs.

Docker is built on the LXC virtualization mechanism. A Docker image is operated in units called containers. Each container uses a process. Depending on its power, thousands of docker containers can work on one machine. Container images share system files that are common. Therefore, disk space is saved. As seen in the figure 3.6, application containers use common bin (exe) and libraries. However, in classical virtual machine systems, separate operating system and library files have to be reserved for each application.

Figure 3.6: Docker Microservices Architecture



Docker takes the image of the installed latest versions of the software and makes them available again. You can create these images once and send them to other servers, or you can create different images on each server. Every server can rebuild the same image by looking at the instruction files called dockerfile. In this way, no manual intervention is required.

Another feature is that Dockerfile and images can be improved. If you want to add a few more steps to the instructions, instead of giving commands from the beginning, it continues from where it left off, which saves time.

3.2.1 Classic Vm Vs Docker

VMs have a full operating system for each employee instance. Docker, on the other hand, uses images that are reduced in size instead of the full operating system and uses the host operating system libraries as shared. However, this makes Docker the system resource-friendly and lowers the isolation level. With these, we can make the following comparisons.

Table 3.2: Comparison on Virtual Machine and Docker

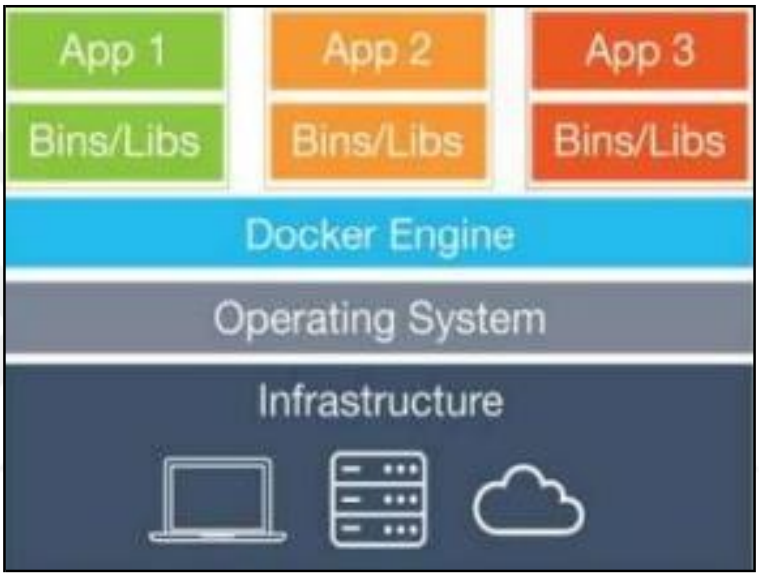
Comparison type	**VM**	**Docker**
OS	Full operating system	Minimized operating system image
Isolation	High	Very Low
Getting it working	Minutes	Seconds
Versioning	Not	High
Easy to share	Low	High

Both virtualization approaches have advantages and disadvantages compared to each other. However, it would be correct to say that some advantages are very critical from the perspective of Docker.

Docker works with hypervisor support, so this platform works virtually and is located between the server and operating systems. Since it is a virtual platform, the server is independent and is used to manage all relevant resources of multiple operating systems on the respective server. Virtualization as Hypervisor is divided into two parts, These are Para-virtualization and Full Virtualization (Felter, 2014). As seen in Figure 3.7,

Docker Container's architecture is fully illustrated. Docker can host and manage windows and linux containers as virtualization. On a single server, it is managed by multiple linux and windows container docker, all of them are isolated from the main server, it only depends on the main server in resource management. In a single container, resources such as Network, Memory, CPU and Block I / O are separated by the linux container, so docker (Seo, 2014). All of these are managed by docker.

Figure 3.7: Docker Service Approach and Architecture



Managing Linux containers is not an easy task as expected, because system resources are involved. Waldspurger has made researches in the sense of managing the CPU of the architectures in linux containers and in the proper management of each of the system resources. In a server, both hyper-v and VMWare virtualization when a structure is presented using its architecture, it is not efficient to use more virtual structures in it (Waldspurger, 2002). This problem of not working efficiently by docker linux containers it is handled again and the problem is solved. Because the containers contain the necessary resources to the applications they contain and they work efficiently in consuming system resources.

Therefore, even if it is a server with low system resources, containers can be created as much as possible. Let's explain this structure by giving examples; Let's say a business

provides email security services. The service provided is to check spam, virus and malware in e-mails. When the product is installed in the cloud, it contains extra features, one of which can transmit daily reports to the agent (Van der Aalst, 2004). Applications like this, the system must be independent. So it has to support every system. For this reason, the docker virtualization approach comes to the fore here, because, as mentioned above, there are containers isolated from the server on which it is running. This approach also minimizes the cost of companies that include test servers, because these servers are not actually needed, because both development and testing processes are hosted by different microservices within the same server. In addition, since containers work on minimized operating systems, applications can run quickly. Therefore, containers are much more efficient at this point than VMs.

In addition to all these good aspects, OpenVZ, its docker and competitor, is undergoing a security perspective and significant improvements are made. As containers consume the same resources, any vulnerability in containers can be it may cause leakage. In this respect, VMs are better.

One of them is easy work. Docker Hypervisor runs in seconds because it doesn't use a fully loaded operating system and runs close to the host system. Another is its predisposition to versioning. One of the most striking features of Docker is its versioning feature. Docker allows us to record different versions of the operating system images it uses. This situation opens the door for sharing the prepared images among users. If we talk about another feature, this is shareability. Operating system images prepared by users or distributors can be sent to central servers and can also be obtained from other central servers by other users.

The operation of containers in a single operating system brings the security problem to mind. Docker brought software solutions to this issue. Applications running in the container cannot see the application in another container unless otherwise specified and in other words, they are isolated.

3.2.2 Docker Container Benefits

Docker Linux containers have improved considerably over the past few years and have become popular in terms of flexibility and efficiency. Sharing images, especially the Docker hub, has greatly increased the popularity of the docker. These advantages are speed, portability, scaling, and Rapid Delivery.

3.2.2.1 Speed

Containers have a fast-running structure because they work with minimized small operating systems (Vase, 2015), which, unlike large operating systems, are only operating systems designed to do a particular job, so container's operation is based on VMs. it happens very quickly. It is possible to run processes such as test, deployment, production in docker, and when these speed comparisons are made, it can be measured that there is a tremendous level of difference.

3.2.2.2 Portability

Because the containers are micro services that docker runs, they are independent of the server they are on, so they can easily be moved to other operating systems because the container does not run on the main server, it runs on the minified operating system in the image, in this feature the container ' It increases the portability of the s as much as possible (Vase, 2015).

3.2.2.3 Scalability

Docker allows scalability in structure, whether the application is on a physical server or in a cloud environment, it is possible to increase or decrease the containers according to the need in a structure that works with docker, so that the number of containers will decrease in the spare time of the server when the server is dense. is an issue to be considered especially in terms of resource management and cost management (Joy, 2015).

3.2.2.4 Rapid Delivery

The docker management also has a standardized structure within the company, especially the person who manages the docker is a devops engineer, software engineers

who manage, develop and test applications in docker are software engineers. When we look at development environments such as test, deployment and production (Vase, 2015), building this structure with a docker is a comfortable job. In this way, everyone will know where to bend and an error-free, perfect development doors of the media will be opened.

3.2.2.5 Density

Since there is no virtualization technology called hypervisor in Docker, the management and use of resources will be so efficient. In this way, multiple containers can be operated easily on a single server. If we talk about the performance of the Docker containers, it can be said that it works better because it contains high density and when looking at resource management points (Joy, 2015).

3.2.3 Disadvantages Of Docker Container

It is an undeniable fact that Docker has disadvantages, unlike its advantages, because every technology has a flaw, and those flaws can be improved over time. There are related items below;

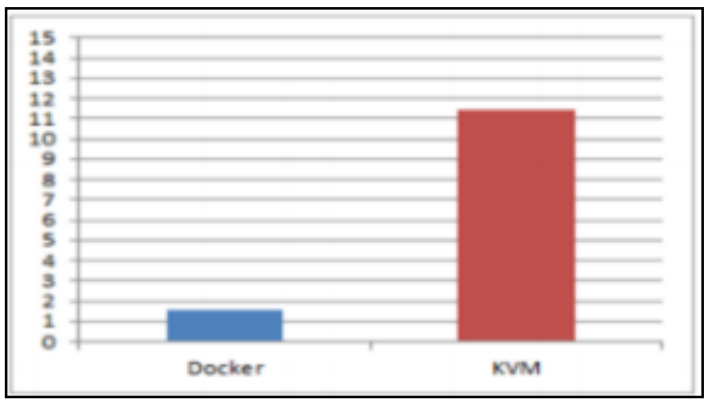
a. Completely virtualization is not possible, because the docker still runs as a service on the server it is on, so any problems that may occur in the server can affect the docker if it wants to.
b. Docker does not work on older machines, so you will need to use a server that supports 64-bit.
c. All virtualized media should be supported with docker containers on all servers. Although the tool called Boot2docker does this job,
d. It should be checked whether the systems are permitted by the administrators or if the server is integrated with the operating system or if there are performance leaks (Harji, 2013).
e. When you use Docker, security issues are of great importance, because when you use virtualization technology, applications, ports, vulnerabilities in all virtualized small containers should be resolved.

3.2.4 Comparison On Docker And Various Technology

Docker virtualization technology also has competitors, and through these competitors, the efficiency of containers will be measured and the efficiency of other container technologies will be measured in the meantime. The containers on which Docker is working as a service directly in cloud technologies are insulated (Seo, 2014). The statistic of spending CPU and other resources is minimal, both for this reason and because containers work on minified operating systems. Especially boot time and image build time are much faster. Such features prevent VMs from being used in the Cloud because they are thanks to the advantages. Two servers with the same resources were used in the cloud environment for measurement. The open stack platform was used, the first for the docker and the other with the classical KVM approach. Ubuntu was chosen as the operating system on the server, and the boot times entered were calculated, images were built on both servers.

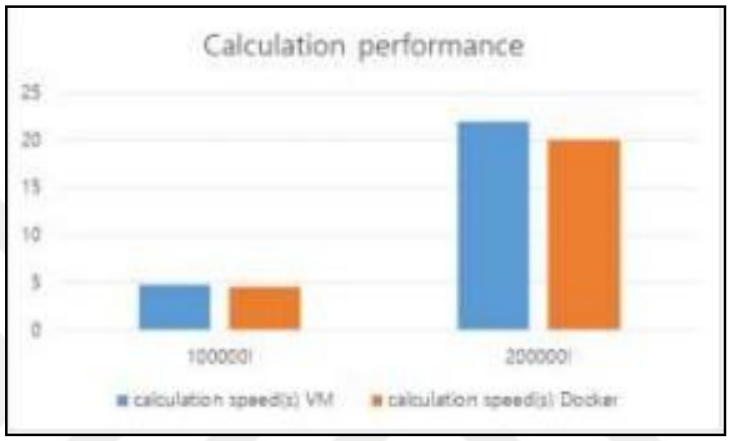
As seen in Figure 3.8, the loading time of the docker is more efficient compared to KVM. Docker works like the KVM guest operating system while using the server resources it is running on. Therefore, when all the above arguments are evaluated, it can be said that the loading speed of the docker is much more efficient than the KVMs.

Figure 3.8: Comparsion on Docker and KVM boot time(Seo, 2014)



Algorithms in python were used to calculate all these processes. As seen in Figure 3.9, containers with a working speed of 100,000 were measured to be around 4.5s. To measure the operating speed, in the docker and KVM, the process was repeated 100 times, and the average processing time and standard deviation were accurately calculated.

Figure 3.9: Docker vs KVM CPU Working Performance (Seo,2014)



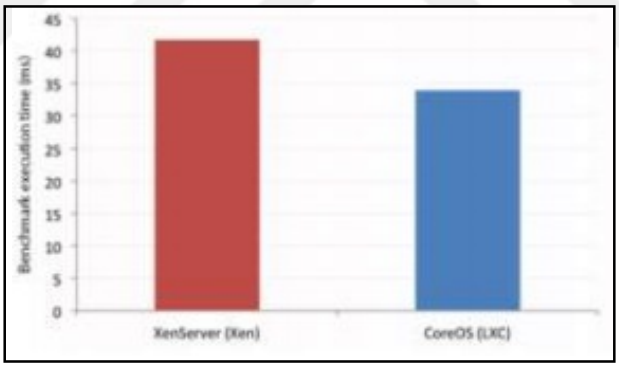
As seen in Figure 3.9, the speed of operation and operation of docker containers is much higher than VM (Scheepers, 2014). Since VM works as a guest operating system, it is completely independent from the operating system it is working on (Seo, 2014). Especially when considering issues such as Network, Security, user and system policy, they are some of the reasons to be chosen. As it is known, Docker does not have a guest operating system, containers do not work specifically on the server it is running on, because containers are known to run on minified OSs. Therefore, it can be easily transferred from one server to another, scaled. When all of these are considered, especially in cloud architectures, issues such as speed and cost reveal how important the docker is.

It is observed that some applications are used in virtualization approaches such as LXC and Xen for comparison within virtualization technology (Sheepers, 2014). Therefore, Docker version was run as 0.11.1 server in Core OS 324.3.0 and XenServer 6.2 and the required results were looked at. Of these systems

Considering its features, 4GB, CPU Intel Xeon Quad core is used and Intel VTX is used as virtualization support. The servers use Ubuntu 12.04 as the operating system, and all containers will operate on both machines. Apache 2.2, WordPress 3.9 and PHP 5.3 will be run on the first virtual machine,

The amount of memory allocated to them is determined as 2GB. This example is an application server example. 1GB of memory is used in the second machine for MySQL database 5.5. This database is actually the database that holds the data of wordpress and the data is filled in that way. The second machine serves as a database server. JMeter was used to compare these two servers. As seen in Figure 3.10, when any select clause is executed, it is seen that less load is used when the LXC is compared with Xen. CPU and Network statistics were kept in this benchmarking process because the comparison is for them.

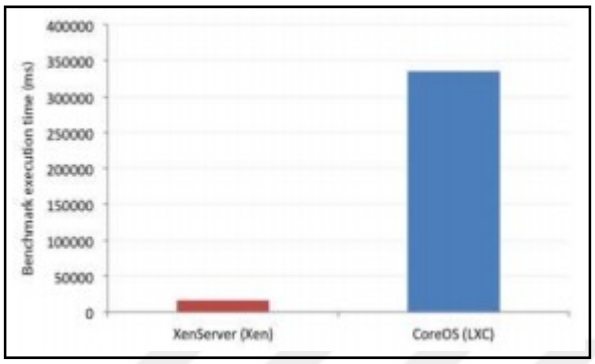
Figure 3.10: Duration one select query between servers (Scheepers, 2014)



As seen in Figure 3.11, it takes 16 seconds on the Xen server to run the insert query in the database, while the LXC takes more time on the installed machine and is approximately

It took 335 seconds. This measurement is proof that LXC containers lack the ability to adequately isolate resources.

Figure 3.11: 10.000 Insert queries durations between servers (Scheepers, 2014)

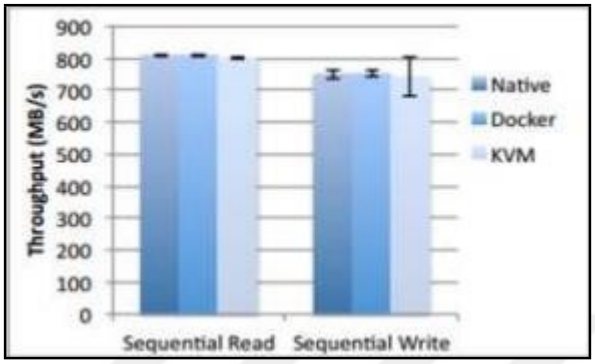


Looking at Xen's sources, it turns out that the sources are not evenly distributed for containers, so this choice was not efficient (Scheepers, 2014). In addition, it has been observed that LXC uses hardware resources efficiently and its isolation is much more successful. Especially in cloud technologies, LXC has been shown to provide superior performance. These virtualization technologies, such as Native, Docker and KVM, have been tested and performance results have been drawn (Felter, 2014).

Approaches that utilize hardware resources completely were investigated. To build all these processes, the IBM x3650 M4 server, 16-core Xen E5-2665 processor and 2.4-3.0 GHZ

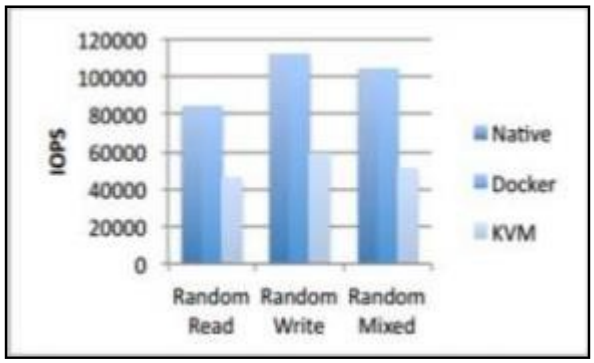
two Intel Sandy Bridge-EP and 256 GB ram are used. To build an inefficient memory access, the two processors are integrated with the QPI connection. It was used as Ubuntu 13.10 as server operating system, 1.0 as docker version, 3.11.0 as Linux kernel,
1.1.1 as libvirt version, and QEMU 1.5.0. As seen in Figure 3.12, sequential reading and writing performances are scaled and each is slightly over 60 seconds, and it has been observed to be an average of 1MB for I / O. Therefore, a very light overhead can be observed by Docker and KVM, whereas in other approaches it has been observed that KVM performs almost four times less.

Figure 3.12: I/O measures sequential process (MB/sec) (Felter,2014)



As seen in Figure 3.13, using SSD technology, irregular I / O operations were performed using 4 kB frame size and 128 concurrency and it was observed how these returned to the server as workload. As it is assumed, although it works much more successfully in KVM I / O operations, there are so many performance leaks, but there is no such measure in the docker.

Figure 3.13. Comparison I/O Random Durations (Felter, 2014)



These technological developments in the field of docker and virtualization in the last 10 years have proven by supporting the criterion of how successful the docker is in these matters.

(Felter, 2014). It is observed that when a deep knowledge container is at the docker point, KVM can perform every operation it can perform more efficiently in itself. Looking at the results, it is observed that there is an overhead for CPU and I / O operations on both the docker and KVM side. In the following processes, different test criteria of these technologies will be illustrated looking back.

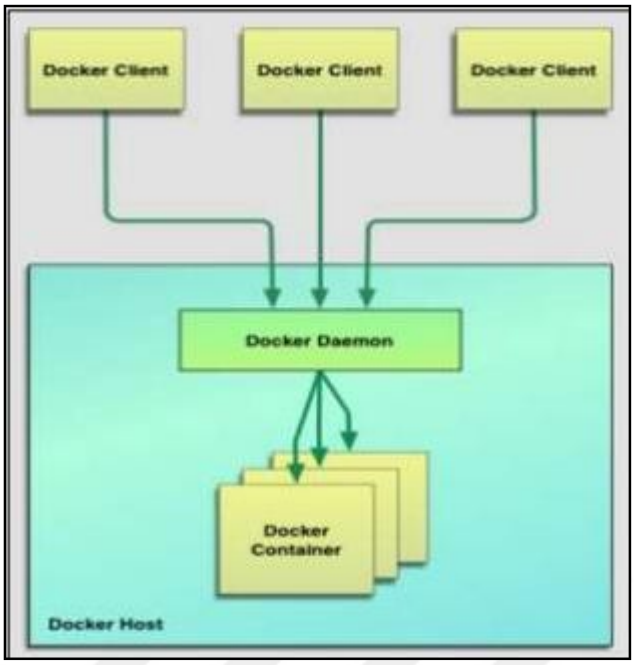
3.2.5 Docker Working Principle

In this section, topics such as how Docker goes a little deeper and works, what it contains, how he does all these processes will be discussed. Details will be given in the docker virtualization and deployment approach used in KVKK application.

3.2.5.1 Docker Machine and Docker Client Side

As shown in Figure 3.14, the docker consists of two structures, the first is the docker daemon server side, and the other is the docker client side where we connect to docker and get information from the docker with API restful requests. The Docker Server can be installed on a server and, if necessary, can be accessed from another server via the server and port, and all structures within the docker can be controlled from this side with the docker client (Turnbull, 2014).

Figure 3.14: Docker approach



3.2.5.2 Docker Images

Image can be created by two methods on the Docker side, one of which is to create an image with a read-only template. Examples of minimized operating system images are these. In these images, images about which operating system will work for the container are created. Ubuntu or Fedore can be given as examples, their versions can be labeled. Tagging can actually be considered as the version of the image, not just the operating system version, but also the version of the entire image that works in it. Another way to create an image is the files called "Dockerfile". To run these files, open CLI in the directory where the file is located and then click "docker build." It is sufficient to run the command, after this command is executed, a series of commands with its own syntax are run in Dockerfile, and the builder of the container is executed, in fact, the image for the container is created here, then the containers can use the relevant images.

3.2.5.3 Docker Registries

Docker registry is actually one of the most important features that make docker images reusable and shared with other developers, it is one of the most important features that makes docker superior because the developers do any specific job images (image can be shown under the sftp service in KVKK application if it is given as an example) can push

or pull over a source. Docker Hub was set up for this, especially developers can share images, shoot images, thereby ensuring sharing.

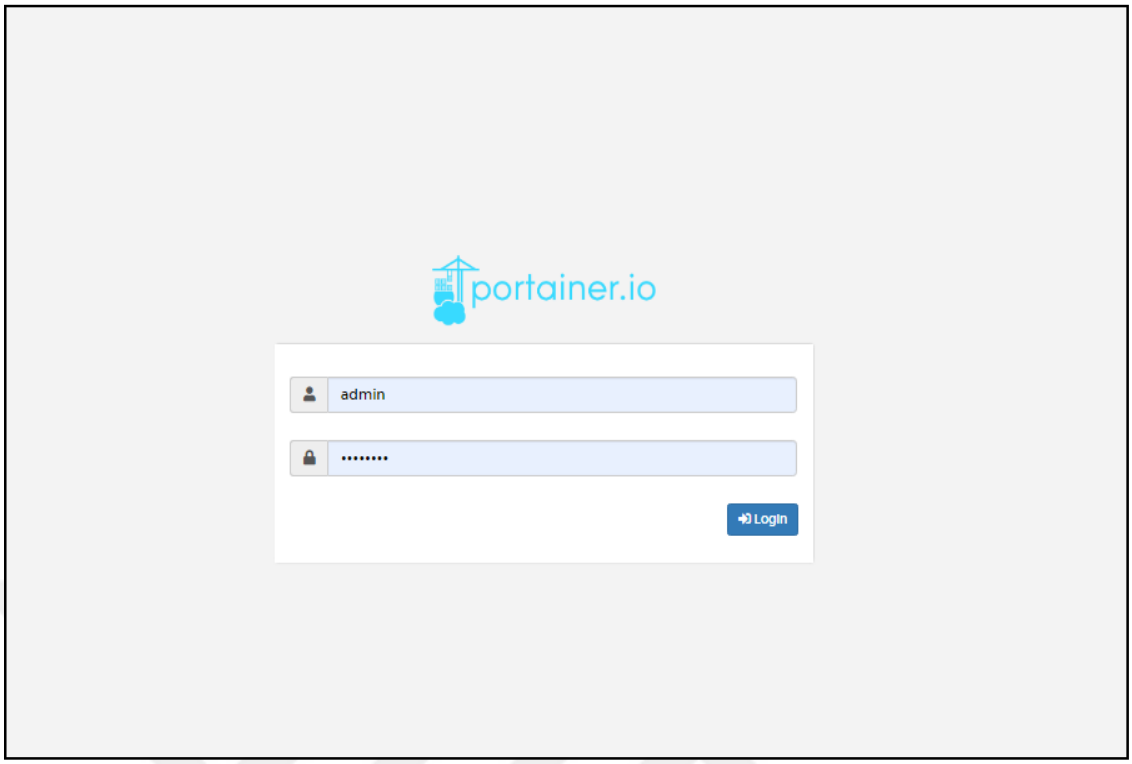
3.2.5.4 Docker Containers

The docker image was mentioned in the title above, the docker container is actually the service itself on the docker side. However, they need an underlying infrastructure, namely docker images. For example, with Sftp Dockerfile, we create an image to perform sftp operations, then use this image to create a container to perform file management operations.

3.2.6 Docker System Approach Used In Application

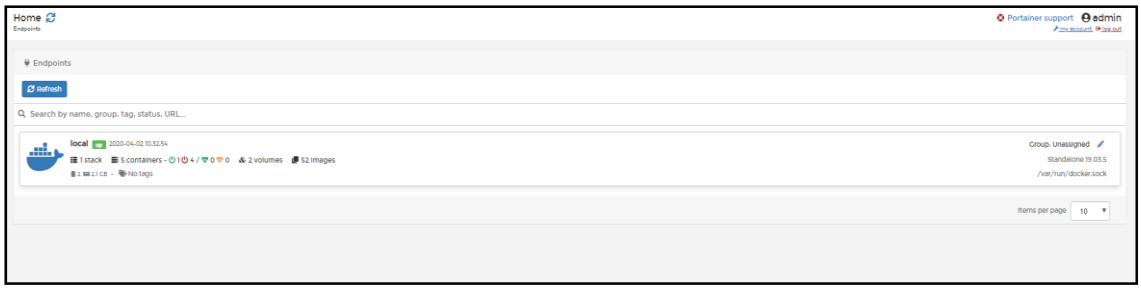
In the system approach, the docker management tool named portainer was used to manage the docker. Portainer is a tool that contains many features and also provides an effective docker management, which is also a docker image, and it is very easy to manage the docker thanks to its web interface. Below is a screenshot for the portainer.

Figure 3.15: Portainer Login Page



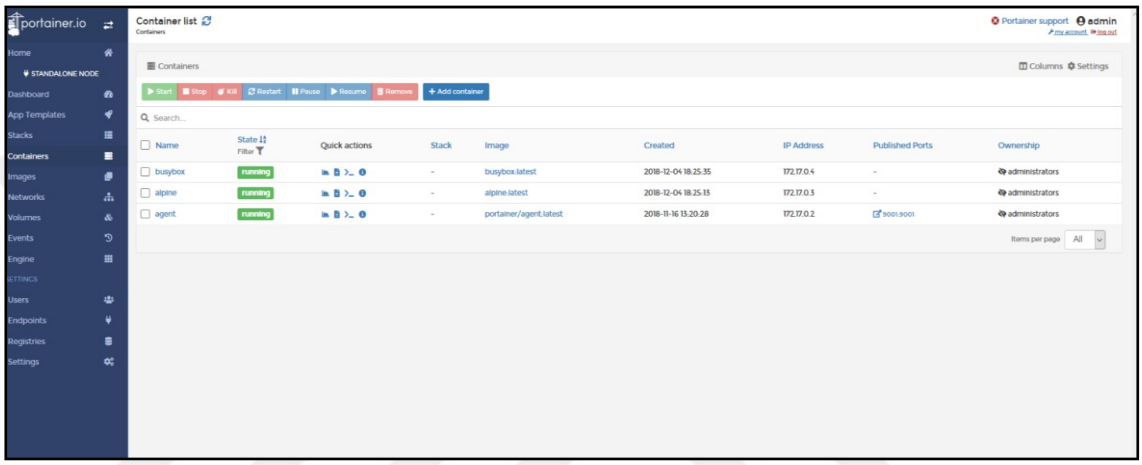
As it is seen, it also provides convenience in the security field, a person who wants to use portainer needs to know the admin password beforehand. In this way, it can manage the docker through the portainer.

Figure 3.16: Portainer Docker Machine List



Here, if there is more than one docker machine in your system, it presents their list, also how many container stacks it contains, how many containers it contains, how many of them are working, how many of them contain hosted information, how many volumes (the files the containers hold permanently) and there is permanently present image information.

Figure 3.17: Portainer Docker Container List



It also has features such as seeing the status of each container in the cluster, examining its logs, and giving command lines to connect to the containers and perform the necessary operations.

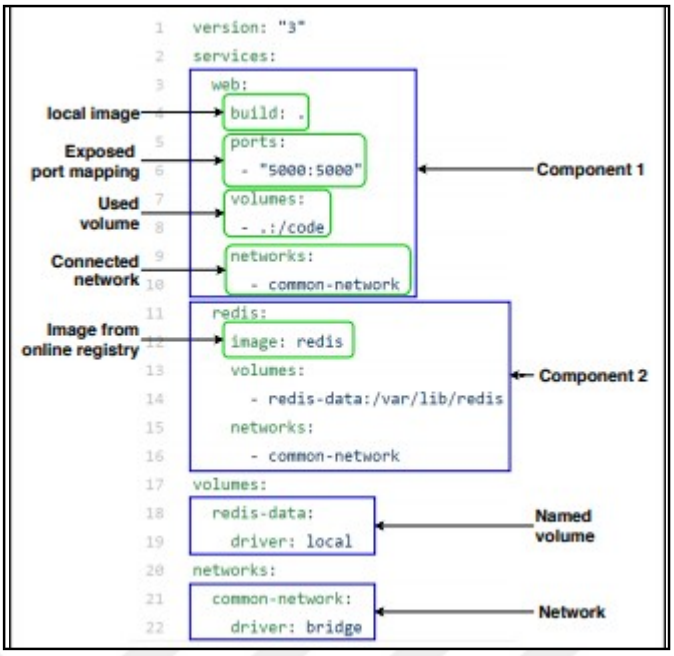
3.2.7 Docker Compose

Docker compose is a tool for identifying systems or applications containing multiple containers. With a single command, it makes it possible to stand up all the services with the configuration in the file. Docker compos is used in KVKK application.

One can compose a multi-component application using Docker Compose which composes a set of components, each of which is an image and a set of options that specify how the component should behave. One can reuse the same image for different components; the reused images will result in different component once instantiated.

Such composition of components is specified using a configuration file such as docker-compose.yml. Figure 3.18 shows an example of a Docker Compose file, which composes a multi-component application from the web and redis components. The webcomponent is represented by a local image (built from a local Dockerfile). Theredis component is created from the "redis" image, which is hosted on online registry (DockerHub, 2019). Additionally, both components (in Figure 3.18) have additional options that configure their environments. The web component exposes the port 5000, stores its data in an external volume (stored in the "/code" path in the host machine), and uses the common-network specification to access other components of the same multi-component application.

Figure 3.18: Components and options in a Docker Compose file



3.2.7.1 Docker Compose Features

The features that make Docker Compose more effective are:

a. It provides multiple isolated environments on a single host,
b. Once the container is created, it is possible to transfer the files in the predefined host to the relevant container,
c. Creating containers again only when there is a change,
d. Variables and moving a composition between environments.

3.2.7.2 Docker Compose File in Application

In our multi-application, there are 4 services in docker-compose file. These are, app-server, app-client, db, SFTP and there are 2 network which will be created when the docker-compose runs.

Figure 3.19: Backend Service in KVKK application.

```
# App backend service
app-server:
  # Configuration for building the docker image for the backend service
  build:
    context: kvkk-vektora-server # Use an image built from the specified dockerfile in the `kvkk-vektora-server` directory.
    dockerfile: Dockerfile
  image: kvkk-vektora-server:latest
  ports:
    - "8443:8443" # Forward the exposed port 8080 on the container to port 8080 on the host machine
  restart: always
  depends_on:
   - db # This service depends on mysql. Start that first.
  environment: # Pass environment variables to the service
    SPRING_DATASOURCE_URL: jdbc:mysql://db:3306/kvkkadmin?useSSL=false&serverTimezone=UTC&useLegacyDatetimeCode=false
    SPRING_DATASOURCE_USERNAME: root
    SPRING_DATASOURCE_PASSWORD: root
  networks: # Networks to join (Services on the same network can communicate with each other using their name)
    - backend
    - frontend
```

There is a image which consists of Dockerfile, when docker-compose builds the app-server service, the image file builds appropriately. In figure 3.20, there is a dockerfile to when app-server builds. There are a lot of steps in file, it inherits from open jdk images and it runs the other parts. First, it adds curl, sets working directory, because subsequent steps will be run in this directory and it copies maven, maven handles to build java source codes and copy pom.xml because when maven runs, some java files have dependencies and they are required to download from Internet, by the way if there is no change in pom.xml, this step omits, it copies the source file to working directory and sets backend application’s working path with docker entrypoint image command.

Figure 3.20: Backend Service Docker Image File

```
#### Stage 1: Build the application
FROM openjdk:8-jdk-alpine as build

RUN apk --no-cache add curl

# Set the current working directory inside the image
WORKDIR /app

# Copy maven executable to the image
COPY mvnw .

COPY .mvn .mvn

# Copy the pom.xml file
COPY pom.xml .

# Build all the dependencies in preparation to go offline.
# This is a separate step so the dependencies will be cached unless
# the pom.xml file has changed.
RUN ./mvnw dependency:go-offline -B

# Copy the project source
COPY src src

# Package the application
RUN ./mvnw package -DskipTests
RUN mkdir -p target/dependency && (cd target/dependency; jar -xf ../*.jar)

#### Stage 2: A minimal docker image with command to run the app
FROM openjdk:8-jre-alpine

ARG DEPENDENCY=/app/target/dependency

# Copy project dependencies from the build stage
COPY --from=build ${DEPENDENCY}/BOOT-INF/lib /app/lib
COPY --from=build ${DEPENDENCY}/META-INF /app/META-INF
COPY --from=build ${DEPENDENCY}/BOOT-INF/classes /app

ENTRYPOINT ["java","-cp","app:app/lib/*","com.vektora.kvkk.KvkkApplication"]
```

In figure 3.20, there is an image step when the docker-compose builds the image, it add tags, in **ports** step, it declares the application output ports, it was chose 8443, because 8443 port is actually tomcat SSL port for increasing the security (Apache Tomcat, 2020). In the **restart** step, it declares the restart always, it means the when the container

stops, it will automatically start again. In the **depends_on** step, this service depends on MySQL so the MySQL service must run first. You can declare the environment variables in the **environment** step, in **network** step, it needs to be declared which networks the service connects.

Figure 3.21: Frontend Service in KVKK Application

```
# Frontend Service
app-client:
  build:
    context: kvkk-vektora-client
    dockerfile: Dockerfile
    args:
      REACT_APP_API_BASE_URL: https://v-kvkk.com:8443/api
  image: kvkk-vektora-client:latest
  ports:
    - "443:443" # Forward the exposed port 80 on the container to port 80 on the host machine
  restart: always
  volumes:
    - ./kvkk-vektora-client/log:/var/log/nginx
  depends_on:
    - app-server
  networks:
    - frontend
```

Apart from backend services, there is a volume step in this service. Volume means, when the container runs, you can mount the files to the container from host.

In figure 3.22, it can be showed the frontend docker image file. It configures the main working directory inside the docker image. This is the base directory used in any further RUN, COPY, and ENTRYPOINT commands.

It copies the package.json as well as the package-lock.json and install the dependencies. This is a separate step so the dependencies will be cached unless changes to one of those two files are made. It serves the React application from Nginx.

Figure 3.22: Frontend Service Docker Image File

```
#### Stage 1: Build the react application
FROM node:12.4.0-alpine as build

# Configure the main working directory inside the docker image.
# This is the base directory used in any further RUN, COPY, and ENTRYPOINT
# commands.
WORKDIR /app

# Copy the package.json as well as the package-lock.json and install
# the dependencies. This is a separate step so the dependencies
# will be cached unless changes to one of those two files
# are made.
COPY package.json package-lock.json ./
RUN npm install

# Copy the main application
COPY . ./

# Arguments
ARG REACT_APP_API_BASE_URL
ENV REACT_APP_API_BASE_URL=${REACT_APP_API_BASE_URL}

# Build the application
RUN npm run build

#### Stage 2: Serve the React application from Nginx
FROM nginx:1.17.0-alpine

# Copy the react build from Stage 1
COPY --from=build /app/build /var/www

# Copy our custom nginx config
COPY nginx.conf /etc/nginx/nginx.conf
COPY ./ssl/certs/v-kvkk_com.crt /etc/ssl/certs/v-kvkk_com.crt
COPY ./ssl/private/server.key /etc/ssl/private/server.key

# Expose port 3000 to the Docker host, so we can access it
# from the outside.
EXPOSE 443

ENTRYPOINT ["nginx","-g","daemon off;"]
```

Another service is database service, this service creates MySQL database client and creates a container for related database operations. This is the container that performs all the tasks of the multi-tenant db approach. This service does not have any written Dockerfile file, it uses MySQL image directly and its tag number is 5.7. In Figure 3.23,

we see a command step, in fact, these command lines are working right after the service occurs, and the command is to determine the character sets of all opened databases and tables using **mysqld**.

In the same way, here it is connected to the backend network as a network, so it is completely isolated as a network by the frontend, and cannot be reached by the frontend, this is done for security.

Figure 3.23: Database Service in KVKK Application

```
# Database Service (Mysql)
db:
 image: mysql:5.7
 command: --default-authentication-plugin=mysql_native_password
 ports:
    - "3360:3306"
 restart: always
 environment:
  MYSQL_DATABASE: kvkk
  MYSQL_USER: root
  MYSQL_PASSWORD: root
  MYSQL_ROOT_PASSWORD: root
 volumes:
  - ./kvkk-vektora-db/data:/var/lib/mysql
  - ./kvkk-vektora-db/dump:/docker-entrypoint-initdb.d
 networks:
    - backend
 command: ['mysqld', '--character-set-server=utf8', '--collation-server=utf8_turkish_ci']
```

Another service in the Multiple Kvkk application is the SFTP service where document management is performed. As seen in Figure 3.24, this service lifts an SFTP server.

SFTP Server is the server where the documents are kept, and the backend can be accessed by SFTP client method, and related document operations are provided in a healthy way. Service contains special scripts, and creates file paths for all customers during the creation of the service, one by one. Atmoz SFTP was used as an image.

Figure 3.24: SFTP Server in KVKK Application

```
# Sftp Server
sftp:
  build:
    context: kvkk-vektora-sftp
    dockerfile: Dockerfile
  image: atmoz/sftp
  volumes:
      - ./kvkk-vektora-sftp/host/upload:/home/kvkk/upload
      - ./kvkk-vektora-sftp/scripts/:/etc/sftp.d/
  ports:
      - "2222:22"
  command: kvkk:123456:100
  networks:
    - backend
```

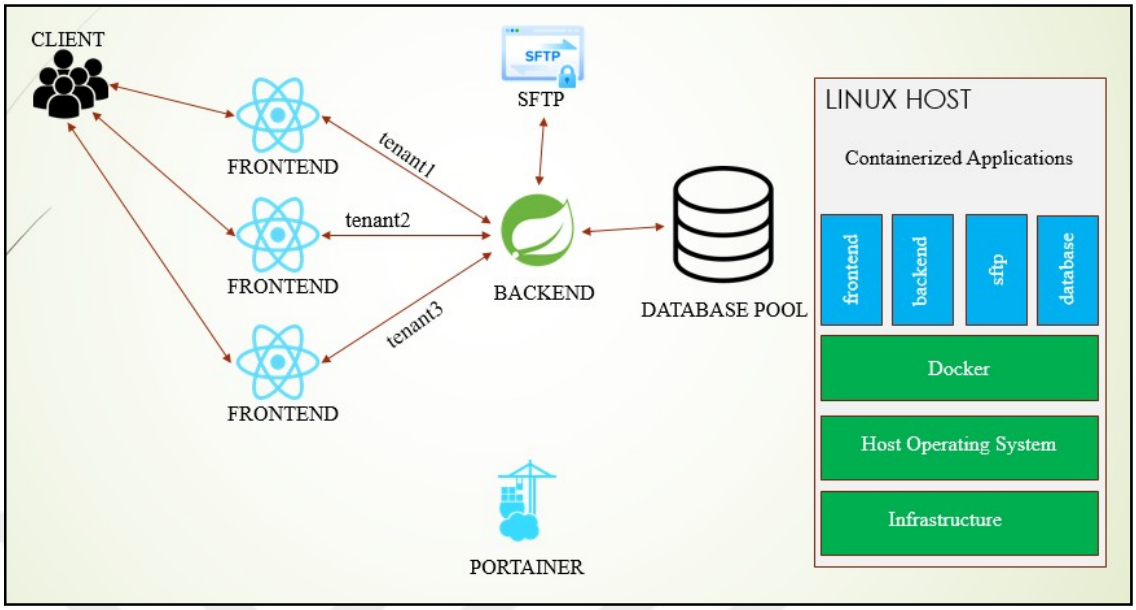
3.2.8 Docker Micro Services Architecture

Docker compose has multi-container execution structure on a single host. As seen in Figure 3.25, the entire architecture is illustrated.

In this architecture, the customer side reaches its own subdomains, then requests the API services with the tenant id in their subdomains, the API services send the related requests either to the SFTP side or to the database side.

By establishing this infrastructure, the architecture in which KVKK processes will be carried out was created. RHEL7 is used as a host, RHEL is a high security Linux operating system used in corporate applications, it operates as a docker service in Linux, and backend, and frontend, db and SFTP are working as micro services in docker.

Figure 3.25: KVKK application Micro service Architecture



3.3 BACKEND APPROACH

JAVA's high level and object oriented software language has been used as the backend approach. The reason why JAVA was preferred is that it has been the one that has been standing in the programming industry for more than 20 years and is the most stable and

Accepted as the language adopted.

(https://www.cloudfoundry.org/wp-content/uploads/Developer-Language-Report_FINAL.pdf)

Over the past nine months, Java and JavaScript remained the two dominant languages. During two rounds of research in the corporate developer environment, industry research firm RedMonk has released the survey that brings JAVA to the top in its corporate web applications. Other languages can be viewed below.

Figure 3.26: Web Application Backend Langıage List

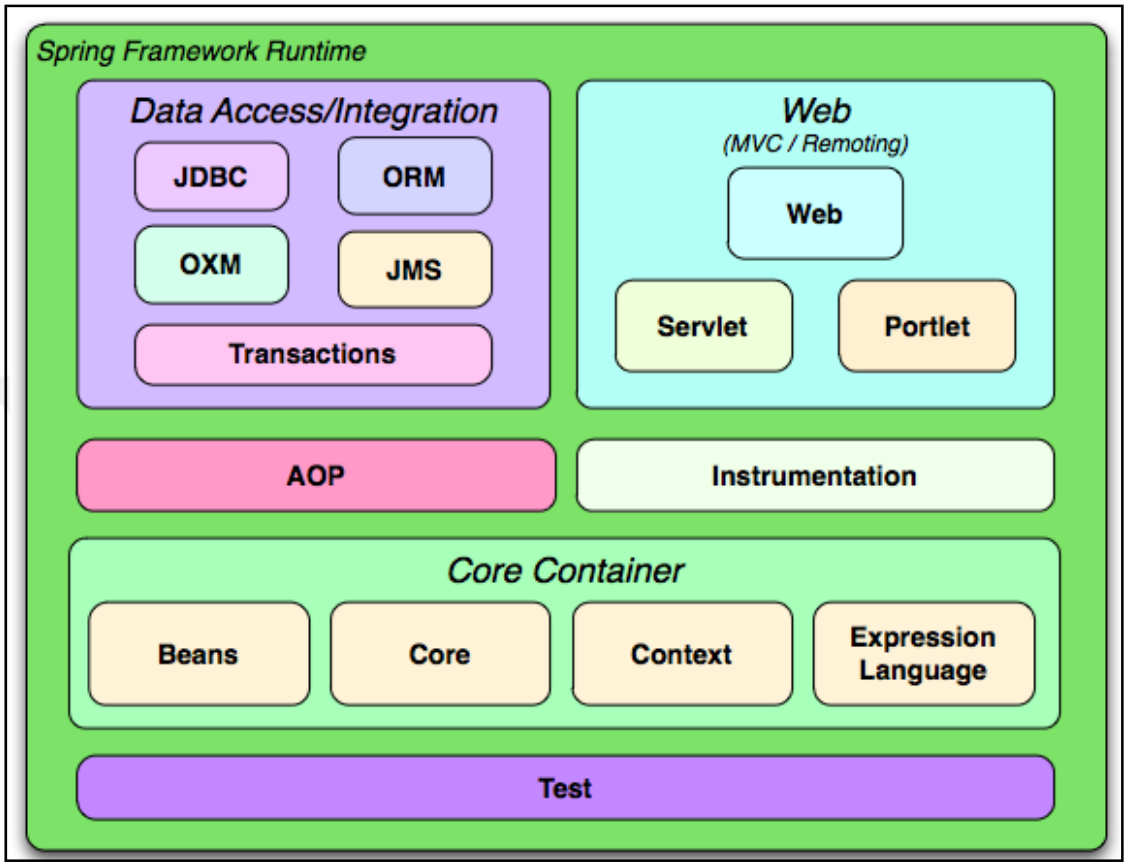
LANGUAGE	NOV 2017	MAR 2018	NOV-MAR
Java	57	58	+1
JavaScript	54	57	+3
C++	45	46	+1
C#	28	26	-2
Python	30	25	-5
PHP	22	22	--
VB.NET	19	17	-2
C	19	16	-3
Visual Basic 6	18	16	-2
VBA	16	15	-1
Perl	16	12	-4
Ruby	9	7	-2
Swift	6	6	--
Typescript	5	5	--
Objective-C	6	5	-1
Assembly	6	4	-2
Matlab	7	4	-3
R	4	3	-1
Scala	5	3	-2
Go	3	2	-1
Groovy	3	2	-1
Haskell	2	2	--
CoffeeScript	2	1	-1
Lua	3	1	-2
Other	3	3	--

3.3.1 Spring Framework

It is an open source framework developed for Spring, Java and .NET. It makes it easier to develop JavaEE applications. We can think of Spring as "frameworks of framework" because it is a framework that supports the use of frameworks such as Struts, Hibernate, Tapestry, EJB, JSF.

Spring consists of approximately 20 different modules divided into groups.

Figure 3.27: Web Application Backend Langıage List



3.3.2 Jwt Security

In the application, JWT security approach has been made to call the services on the backend side.

JWT is an open standard that defines a compact and independent way of securely transmitting information between the parties (frontend, backend) as a JSON object. (RFC 7519 https://tools.ietf.org/html/rfc7519). Since this information is digitally signed, it can be verified and reliable. JWTs can be signed using a public / private key pair using RSA or ECDSA with an HMAC algorithm.

Although JWTs can be encrypted to provide privacy between the parties, we will focus on signed tokens. Signed tokens verify the integrity of the prompts it contains, while encrypted tokens hide these requests from other parties. When the tokens are signed using public / private key pairs, the signature will only be used by the party holding the private key also confirms that it is the signing party.

3.3.2.1 What Is the JSON Web Token Structure?

In its compact form, JSON Web Tokens consist of three parts separated by dots (.),

Which are:

a. Header,
b. Payload,
c. Signature.

The header typically consists of two parts: the type of the token, which is JWT, and the signing algorithm being used, such as HMAC SHA256 or RSA. The second part of the token is the payload, which contains the claims. Claims are statements about an entity (typically, the user) and additional data. To create the signature part you have to take the encoded header, the encoded payload, a secret, the algorithm specified in the header, and sign that.

Therefore, a JWT typically looks like the following.

“xxxxx.yyyyy.zzzzz”.

Header: The header typically consists of two parts: the type of the token, which is JWT, and the signing algorithm being used, such as HMAC SHA256 or RSA. For example:

Figure 3.28: JWT Header Part

Many different algorithms such as HS256, HMAC SHA256 or RSA can be used in the algorithm section. In the Type section, JWT is written. This part is encoded with Base64 and forms the first part of the token to be created.

Then, **payload JSON** is Base64Url encoded to form the first part of the JWT. The second part of the token is the payload, which contains the claims. . Claims are statements about an entity (typically, the user) and additional data. There are three types of claims: registered, public, and private claims. We use the registered claims.

Registered claims are a set of predefined claims which are not mandatory but recommended, to provide a set of useful, interoperable claims. Some of them are: iss (issuer), exp (expiration time), sub (subject), aud (audience), and others.

Putting all together, the output is three Base64-URL strings separated by dots that can be easily passed in HTML and HTTP environments, while being more compact when compared to XML-based standards such as SAML.

Figure 3.29: JWT Full Token Part

3.3.2.2 JWT Token Generator Method

We sign our back-end services with the HS512 algorithm and also with the secret key we have created. Token expiration time is 30 minutes. Let's try a little, let's ask for a new token from the login service. We have obtained the token as seen in figure 3.30.

As you see, signature algorithm is HS512, I choose HS512, because it is more powerful and secure encryption algorithm, As claims, I choose issue date, token expiration time and sign with HS512 and JWT Secret key. When all this is combined, the communication between the frontend and the backend is strong and secure.

Figure 3.30: JWT Token Generator Method

```
@Value("${app.jwtSecret}")
private String jwtSecret;

@Value("1800000 ")
private int jwtExpirationInMs;

public String generateToken(Authentication authentication) {

    UserPrincipal userPrincipal = (UserPrincipal) authentication.getPrincipal();

    Date now = new Date();
    Date expiryDate = new Date(now.getTime() + jwtExpirationInMs);

    return Jwts.builder()
            .setSubject(Integer.toString(userPrincipal.getId()))
            .setIssuedAt(new Date())
            .setExpiration(expiryDate)
            .signWith(SignatureAlgorithm.HS512, jwtSecret)
            .compact();
}
```

3.3.2.3 When Should JSON Web Tokens Be Used?

Some scenarios where JSON Web Tokens are useful:

Authorization is the most common scenario for using JWT. After the user logs in, each subsequent request will include JWT and allow the user to access the routes, services

and resources allowed with this token. Single Sign-On is widely used today due to its small overhead and easy use in different areas and is a feature that uses.

Information Exchange JSON Web Tokens are a good way to securely transmit information between parties. Since JWTs (for example, using public / private key pairs) can be signed, you can be sure that the senders are the person they say. Also, since the signature is calculated using the title and payload, the content has not been altered and you can verify.

3.4 FRONTEND REACT FRAMEWORK

We use react as the frontend technology in the KVKK application because react is a powerful JavaScript framework supported by the Facebook community.

3.4.1 Libraries And Frameworks In Frontend

The world of Frontend has improved a lot in recent years, especially with the development of javascript and the discovery of the V8 engine, many frameworks and libraries have evolved. All excellent features in their field These approaches, which hosted them, also had their equivalents in the github world, especially a community of each framework and library has been created, where numerous bugs have been resolved and new features continue to be added. In Github, you can see how much these approaches are used, how much they are downloaded and usage statistics. There is a table below that explains which framework Frontend developers especially prefer.

Table 3.3: Framework and Library Statistics in GitHub on May 2018

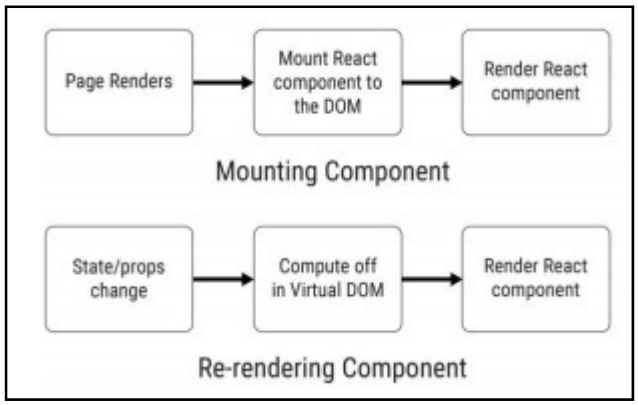
	Angular 1	Angular 2	React	Vue
Downloads	1.4 Million	2.6 Million	9.2 Million	1.5 Million

As shown in Table 3.3, React is the most popular among them, because the react framework was written by facebook, react, which was not successful at first and whose source code is closed, today it continues as an open source. Angular comes right after, Angular is available in versions 1 and 2, in addition there is Vue.js, there are dozens of frameworks that are not listed, but we will get the most comprehensive among them. The React framework was chosen by comparing these statistics in the KVKK application.

3.4.2 Difference Between React.Js And React Native

Offering an excellent user experience, brands such as Facebook and Instagram developed the React Javascript library to develop the website (Gackenheimer, 2015). After React reached a certain maturity, Facebook opened it as open source code to global developers and companies, Javascript ES6 lies on the basis of React.js (Gackenheimer, 2015). In addition, Facebook has derived an additional library called React native from React.js for the compatibility of React on mobile devices and the release of applications close to native, making it possible to make mobile applications close to native with these libraries (Eisenman, 2015) .

Figure 3.31: React Component Mounting Process (Eisenman, 2015)



As shown in Figure 3.31, the concept of mounting underlying React technology is shown, Mounting means whether a component will be on the page. React means any content on its website, When connecting to the Data Object Model, it takes a different approach. With Chrome's V8 technology, deriving HTML from Javascript has become much faster, so instead of dynamically writing all pages in HTML, It is possible to create dynamically with the web language Javascript, because Javascript's code structure enables us to follow a much more effective way in this sense, as it is a powerful community that supports algorithms such as writing algorithms and HTML derivation quickly. React creates a virtual DOM before importing a component to the DOM (Gackenheimer, 2015). In sites made with general HTML, the user needs to refresh the page completely when the data is refreshed or when the sub-pages are visited, that is, the DOM is re-created. This means requesting the files again on the Server side, this way allows the user to makes it difficult and spends a lot of resources. Unlike the traditional method, React manages such changes in the virtual DOM, that is, it compares an updated virtual DOM with the old DOM, if there are changes, the changes are transferred to the real DOM. In this way, there is no need to go to the server again and again, websites become faster.

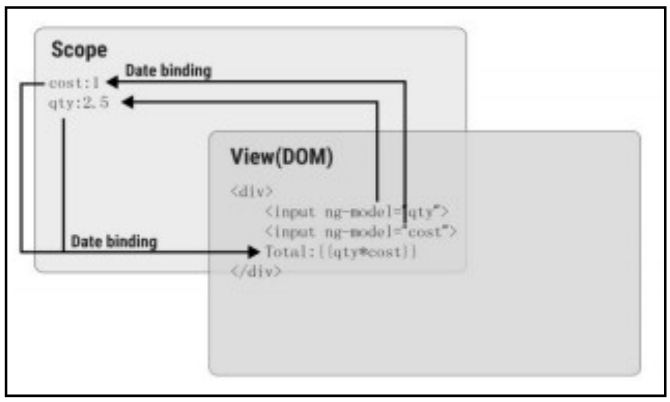
The main advantage of using virtual DOM is that the website becomes much more effective and faster with React. In addition, React is a modular framework, which means that React is an open source, so all developers have offered all kinds of solutions in projects to handle the challenges they face while doing projects, and suggested it to

react, and React has become more powerful by incorporating some of these enhancements, what is actually said is a team game. Therefore, UI enhancements have increased User Interface Modularity. As an example, a user operates in a text form with the same function but on two different pages, the website is re-created between the user page transitions, and the text forms are re-created. However, since there is modularity to this side in the React approach, that component will not regenerate. In short, React does not only deal with frontend enhancements, but also in user interface design and development takes a new approach.

3.4.3 Angular 1 And Angular 2

Angular, like React, is a frontend framework and was developed by Google in 2010 (Ramos, 2018). Angular was not originally a framework that completely destroyed the web paradigm, it was originally developed to create a web form (Ramos, 2018). With the spread of Frontend technologies, while frameworks are moving forward, Angular has grown slowly and steadily and is designed to develop difficult web applications. However, Angular contained a lot of limitations in version 1 and had a lot of handicap, so Google rewrote Angular in 2016 and released the version of Angular 2 (Chen, 2018).

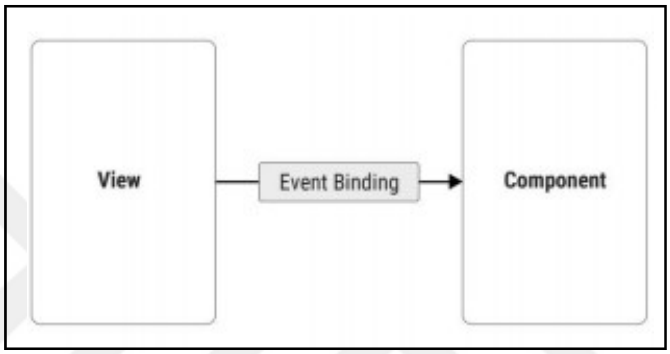
Figure 3.32: Angular 1 Architecture (El Omari, 2018)



The basic approach of Angular 1 was actually based on two-way data binding in web browsers, in fact, this structure is a structure that takes charge and responsibility on the server side (Ramos, 2018). In Figure 3.32 as can be seen, the data binding approach of

Angular 1 is summarized. Here, the part in data binding uses JSON. It is the definition of the whole scope in the entrance and exit sections of the website pages in Angular 1 as a model. Instructions are used to connect to the model (Ramos, 2018). Inputs received from users will be kept in these models and transferred to the server side. Finally, these enhancements are not held on the HTML side and give the pages excellent speed and performance.

Figure 3.33: Angular 2 binding approach



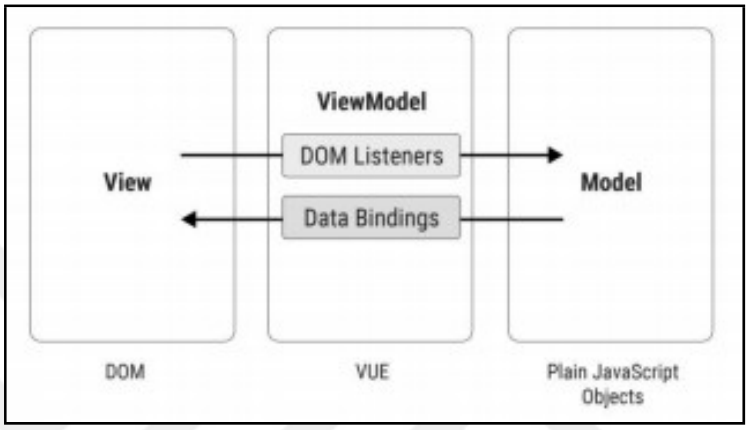
Angular 2 has been rewritten by Google due to the failures of Angular 1, data linking processes in Angular 1 have been optimized. The first is Angular 2, the template directives and controllers was to merge at one point, its name was called Component. As seen in Figure 3.33, Angular 2's event binding has a unidirectional structure unlike Angular 1 (Angular Official Blog, 2018). Another approach is Angular 2, developed with TypeScript developed by Javascript based microsoft (Angular Official Blog, 2018). Fourth, it uses zone.js instead of Scope to track events that interact with each other. Angular 1's goal was to show web apps on mobile platforms too, but this was not very successful, however, in Angular 2 these features have been revised, optimized and have the smaller size and higher speed required for mobile development.

3.4.4 Vue.Js

Vue.js was developed by a developer named Evan You in 2014, it became a frontend framework using Javascript ES6 language (Pastushenko, 2017). The first development purpose of Vue is that it has a structure that can easily communicate with Api and help

design user interfaces. Vue was not usually used in industrial applications, but developers from all over the world helped Vue upgrade, further developing the single- page application example and reaching the desired level. Today, it has become possible to use Vue in all kinds of web applications.

Figure 3.34: Vue Data Binding Approach (Vue Official Blog, 2018)



As seen in Figure 3.34, Vue.js combines with data-driven view, view model, and 3 builds containing the Model. When the view model is examined, this is actually the DOM structure that is the content of the website. DOM listeners are available on the side that displays the model, and structures are available that link data from the user. When users enter data in views (Nikulchev, 2018), my Vue model section monitors the data and uses DOM listeners to update it. It uses the DOM connector to update the website view. In addition, Vue has a two-way structure.

3.5 ANALYSIS OF THE FRONT-END SOLUTION

Taking into account all this information, React, Vue and Angular are groundbreaking frameworks in web technology. In addition, each has different approaches and solutions. Taking this information into consideration, it is necessary to start developing the web in a web page to be developed in order to find out which one is more useful.

3.5.1 Data Processing

In a web application, the ability and wisdom of processing data from users constitute a very important place in terms of the quality of web applications. It is one of the most important parts of Frontend applications.

Table 3.4: Front-end frameworks data binding concepts

	Angular 1	**Angular 2**	**React**	**Vue**
Data Binding	Two-way	One-way/Two-way	One-way	One-way/Two-way

Table 3.4 shows the data binding way in each framework and library. Vue can use DOM listeners to achieve two-way binding or use one-way binding without DOM listeners. The most significant difference between React and Angular 1 is selecting one- way or two-way data binding. Compared to Angular 1, React needs to process more complicated data flow due to check the difference between virtual DOM and displayed DOM. However, using twoway method between component and view may result in component transitionally into unexpected states because of conflicting data being spread from multi sources while one-way binding can avoid the conflict issues in multi data sources especially in event-based circumstance. Therefore, Angular 2 team upgrades the concept that developers can use both one-way and two-way binding. Choose the hybrid method in different circumstances is the best solution to absorb one-way and twoway's advantages.

3.5.2 Volume and Performance

Larger volume means the framework or library contains more features and functions, but it will spend more time on loading the framework or library.

Table 3.5: Front-end frameworks and library memory allocation performance (Stefan Krause, 2018)

Activity	Angular 2	React	Vue
Ready memory	4.8±0.0 (1.4)	4.5±0.1 (1.3)	3.8±0.0 (1.1)
Run memory	10.9±0.1 (2.7)	9.7±0.1 (2.6)	7.5±0.1 (1.9)

Angular 2 has the most significant volume of 143 kilobyte(KB) while Vue has 23KB and React is 43KB. The enormous volume means that Angular 2 has more advanced comprehensive functions and features. However, complicated structure leads to a potential risk that Angular 2's running performance is worse than React or Vue especially in memory allocation. Table 3.5 shows that Angular 2 requires more time to prepare and run the memory while Vue spends less time due to its flexible and efficient advantages.

3.5.3 Language Based

Language-based is another important considerate aspect because different languages have different circumstances in developing the project such as difficulty and efficiency.

Table 3.6: Language-based in front-end frameworks and library

	Angular 1	Angular 2	React	Vue
Language-based	JavaScript ES5	Type Script	JavaScript ES6	JavaScript ES6

Table 3.6 shows the current language-based situation in different frameworks and library. React and Vue use JavaScript ES6 as their based language which is the newest JavaScript industry standard from 2015. Angular 1 uses the JavaScript ES5 which is the previous version. Angular 2 relies on TypeScript provides an excellent experience for type inference and reduces all sorts of bugs from web applications. Furthermore, TypeScript assists developers to get rid of traditional JavaScript programming format with optimizing language structure. However, TypeScript has limited user communities.

There is a potential risk that TypeScript may disappear due to another new syntactical strict superset of JavaScript appear. JavaScript ES6 can be updated but not disappear because it is the industry criterion in developing JavaScript.

3.5.4 Technical Support

Technical support is also necessary because better support can build up a better relationship with developer communities which is an essential part to expand the framework's reputation. React provides strong technical support and its API is very stable. Upgrade and immigration are very simple due to the official scripts to assist developers to finalize the related update. In short, React provides persistent technical services. Angular also provides similar services, but the API is not so durable as React. Some previous version's APIs have been banished. Although Vue has simple immigration and update between different versions, there is no expected update plan from the official team due to the financial limitation.

3.6 KVKK FRONTEND SOLUTION

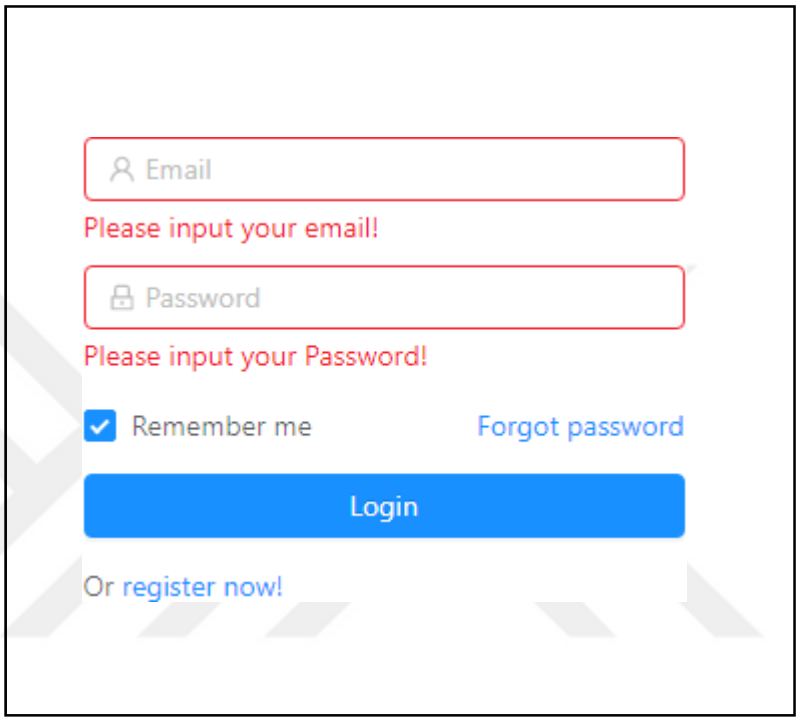
When all the frontend analyzes are considered in our application, React, which is the most suitable framework, has been chosen. Because react fully meets application requests both within the scope of security and with its superiority to its competitors.

3.6.1 Login Page

In computer security, logging in (or logging on, signing in, or signing on) is the process by which an individual gains access to a computer system by identifying and authenticating themselves. The user credentials are typically some form of "username" and a matching "password", and these credentials themselves are sometimes referred to as a login (or a logon or a sign-in or a sign-on). In practice, modern secure systems often require a second factor such as email or SMS confirmation for extra security.

When access is no longer needed, the user can **log out** (**log off**, **sign out** or **sign off**). We also have a fully secure login page in our application. As can be seen in Figure 3.35, email validation is available.

Figure 3.35: Application Login Page.

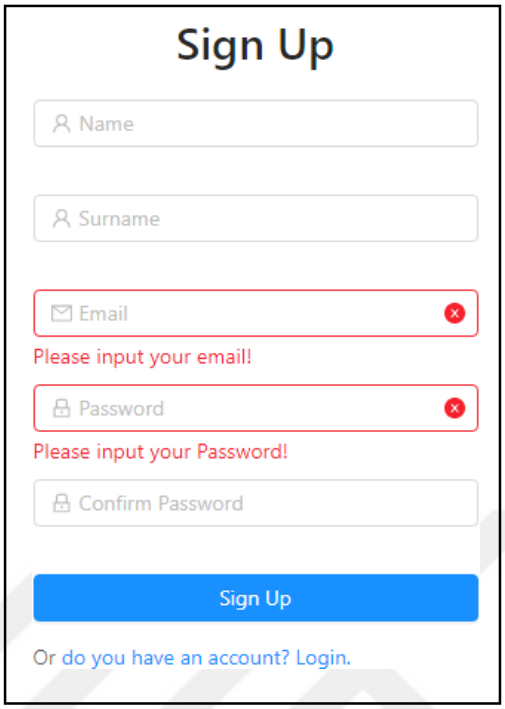


3.6.2 Sign Up Page

Signup pages, also known as "registration pages," aren't all the same. Some are PPC post-click landing pages; others are linked to the homepage via navigation – some even are the homepage. Despite their differences, they all share the same goal: generate sign ups for a service.

In many cases, the signup page is the last step in a business's conversion funnel. It's where prospects navigate after they've evaluated a brand and decided its service offers what they need.

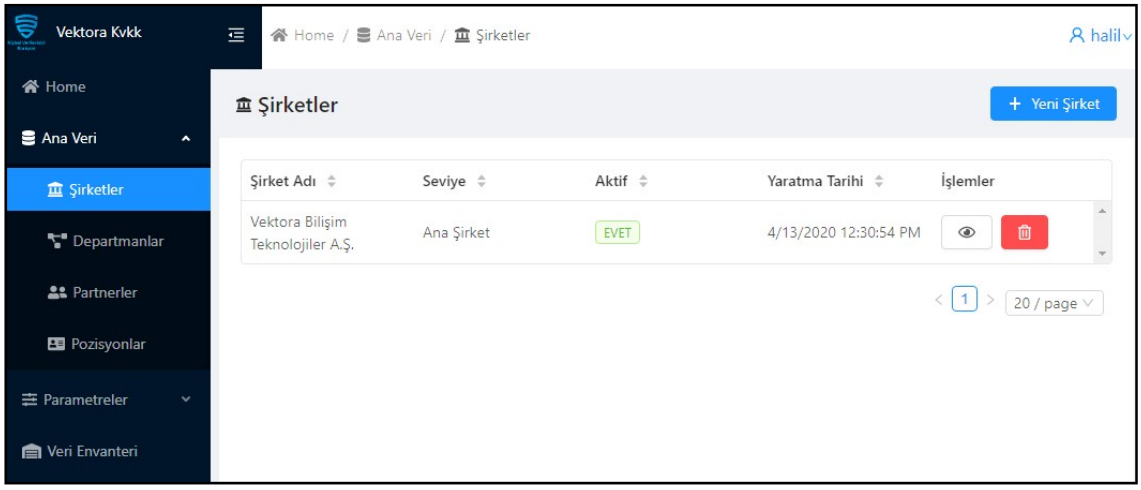
Figure 3.36: Application Sign up Page



3.6.3 Company Page

A company may consist of more than one sub-company, a screen where data entry is required due to a table where all of these companies are kept.

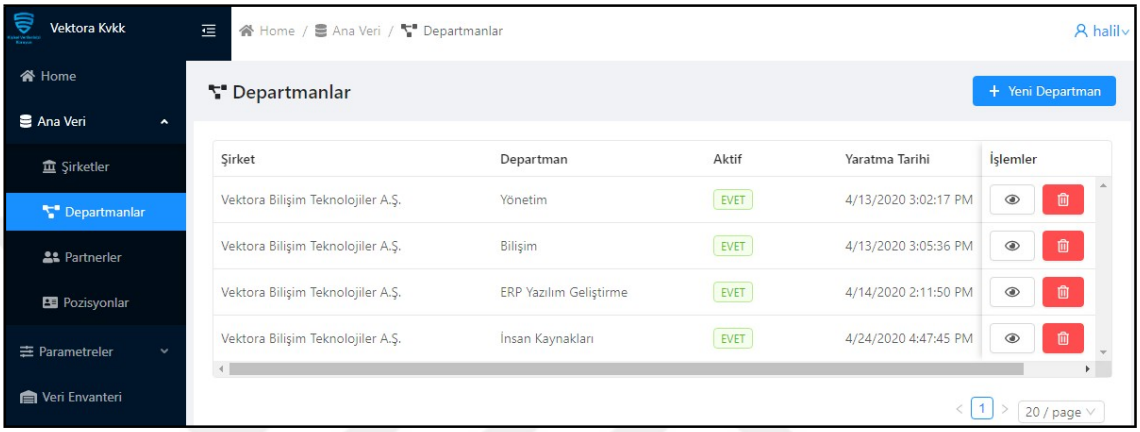
Figure 3.37: Application Company Page



3.6.4 Department Page

Likewise, it has become a must on a screen where people who are subject to KVKK processes can enter and manage department master data with both manual and integration.

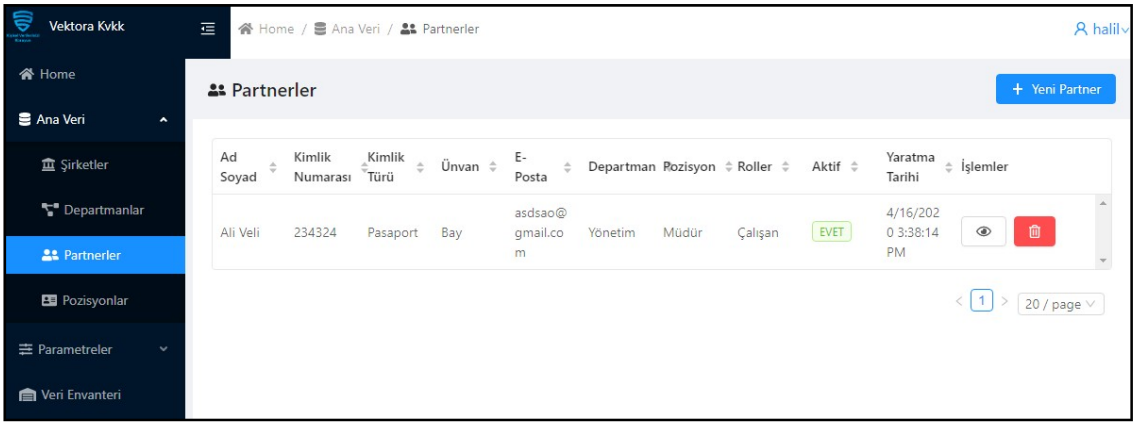
Figure 3.38: Application Department Page



3.6.5 Partner Page

Here, the information of all persons who are subject to the KVKK processes and whose personal data are shared within the scope of KVKK is kept, especially the email field will be an important data in requesting open consent.

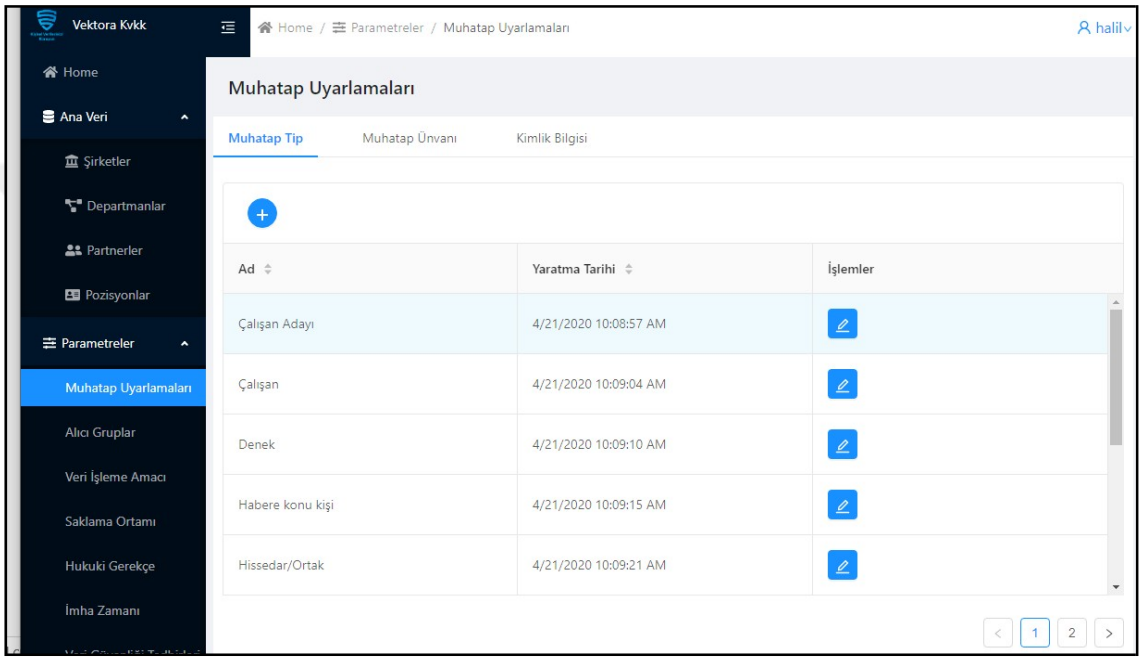
Figure 3.39: Application Partner Page



3.6.6 Interlocutor Adaptations

The type of persons subject to the KVKK process, for example: visitor, customer, employee, and seller, title as an example: male, female, identity information such as: ID card, passport, etc.

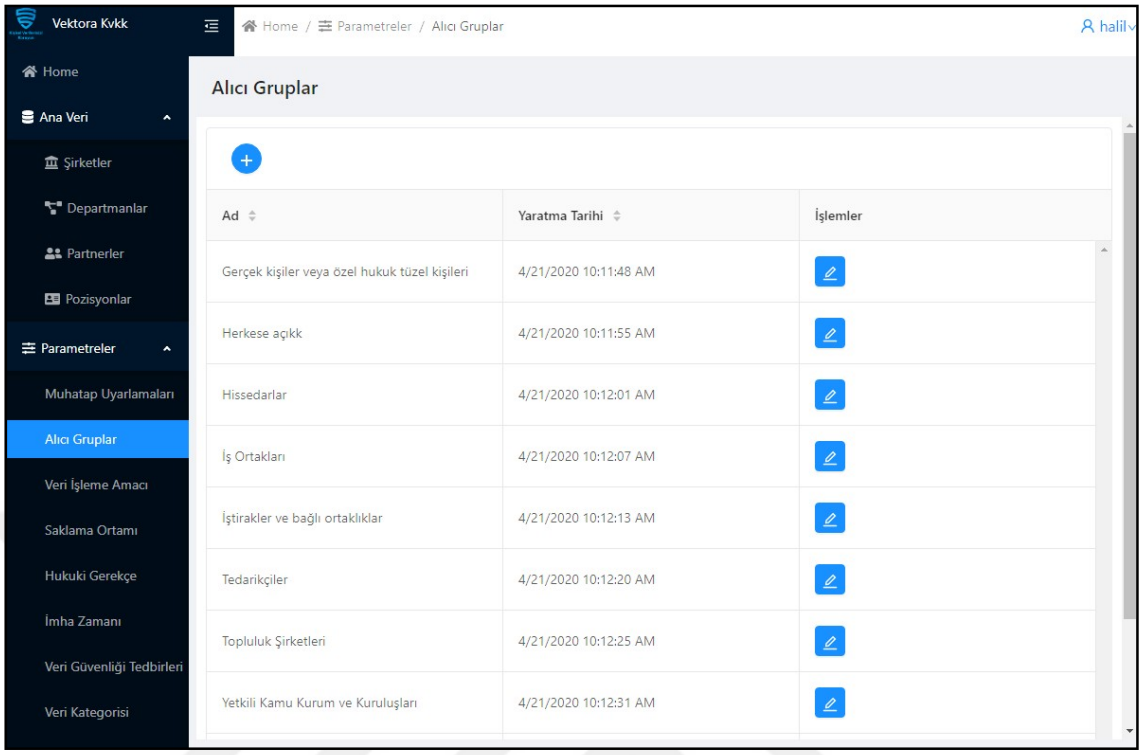
Figure 3.40: Application Interlocutor Adaptations Page



3.6.7 Receiver Groups

When a personal data of the relevant person is shared, it is the master data of the field where the personal data is shared with whom you can see, you can see the samples on the screen. The company must enter the group of recipients when sharing personal data, which corresponds exactly to this field.

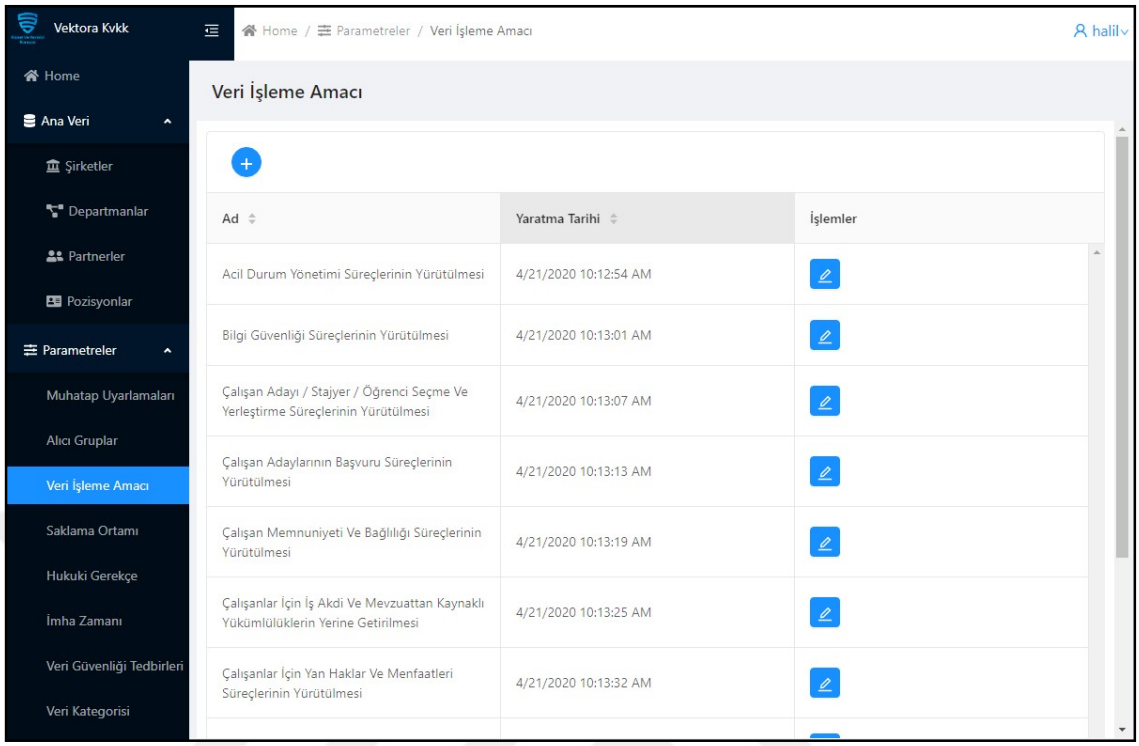
Figure 3.41: Application Receiver Groups Page



3.6.8 Data Processing Purpose

The acquisition, recording, storage, preservation, modification, reorganization, disclosure, transfer, acquisition, availability, classification, or use of personal data by means of wholly or partially automated or non-automated provided that it is part of any data recording system. Anything performed on data such as blocking is a kind of process. For this reason, data processing screen was needed in the application.

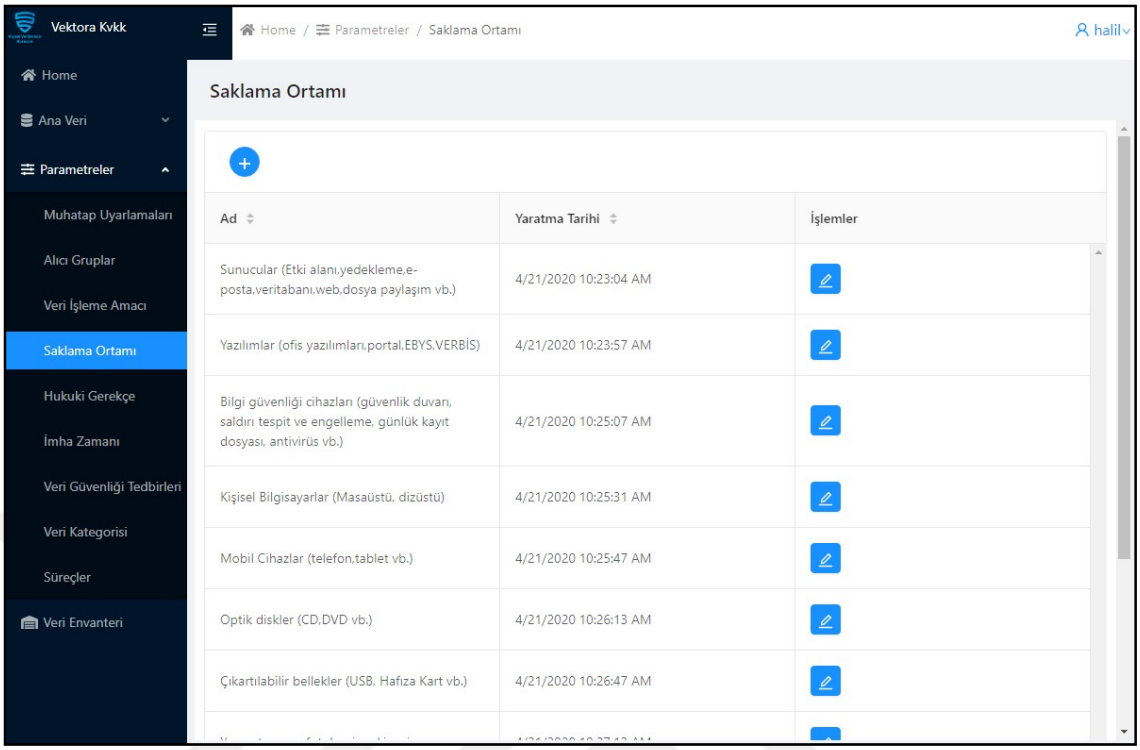
Figure 3.42: Application Data Processing Page



3.6.9 Storage Environment

Environments where personal data will be stored must be entered in the application, for example: servers, software, personal computers.

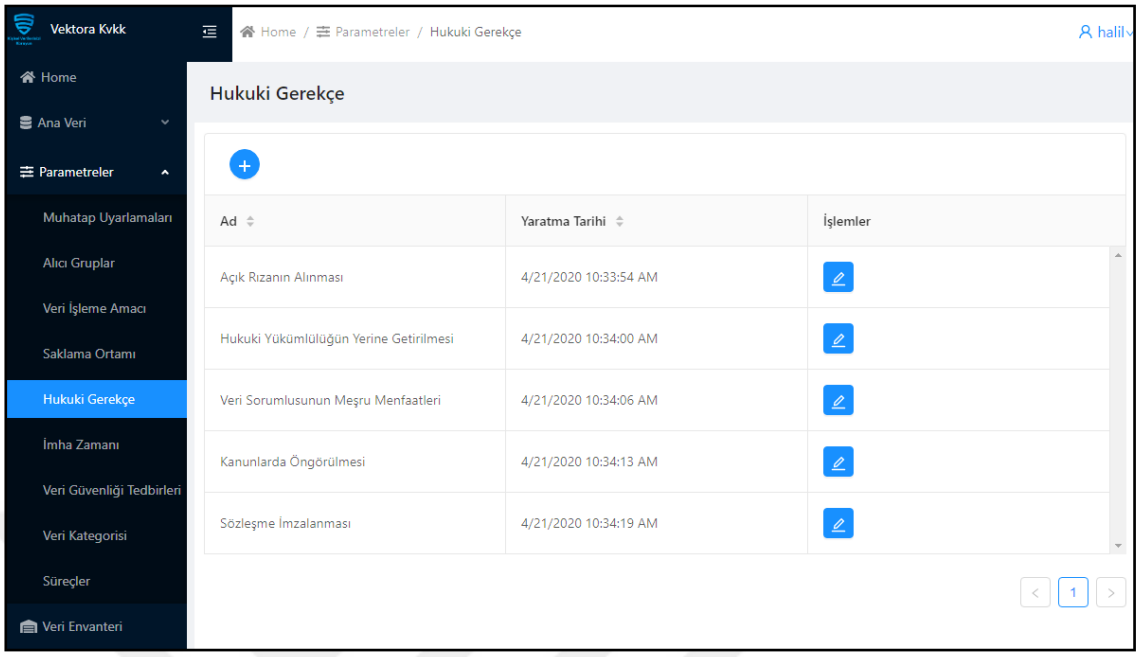
Figure 3.43: Application Storage Environment Page



3.6.10 Legal Justification

The data officer is required to base a legal justification according to the KVKK law when using the relevant personal data.

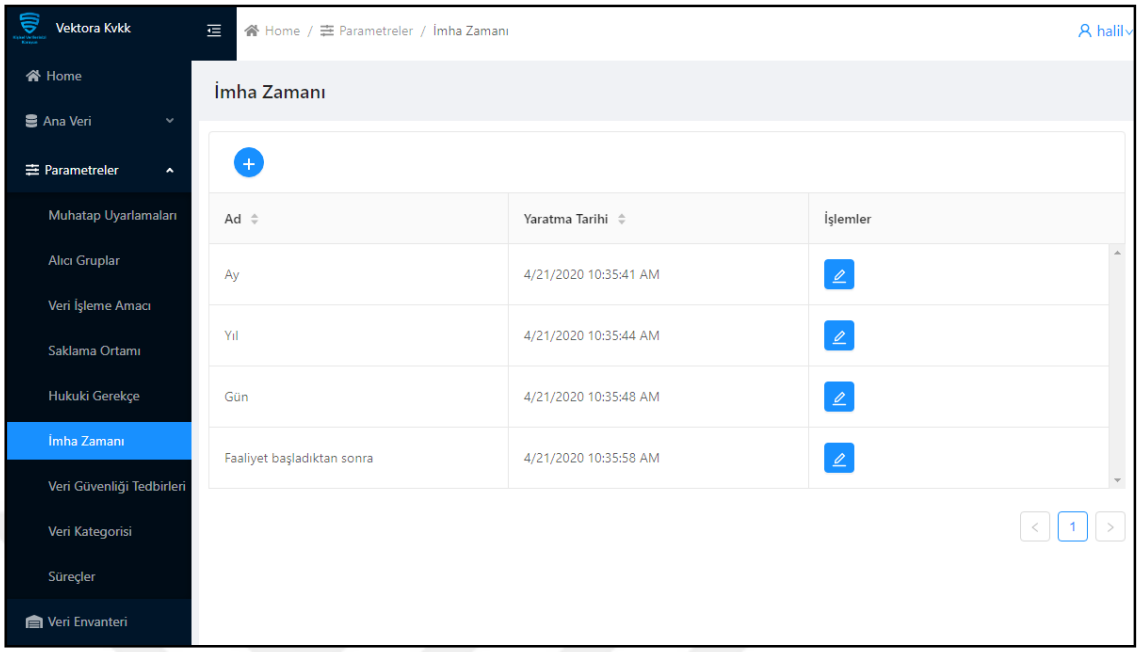
Figure 3.44: Application Legal Justification Page



3.6.11 Destruction Time

Personal data has a duration, or has a renewal period, when it is time, personal data must be retrieved with the consent of those concerned.

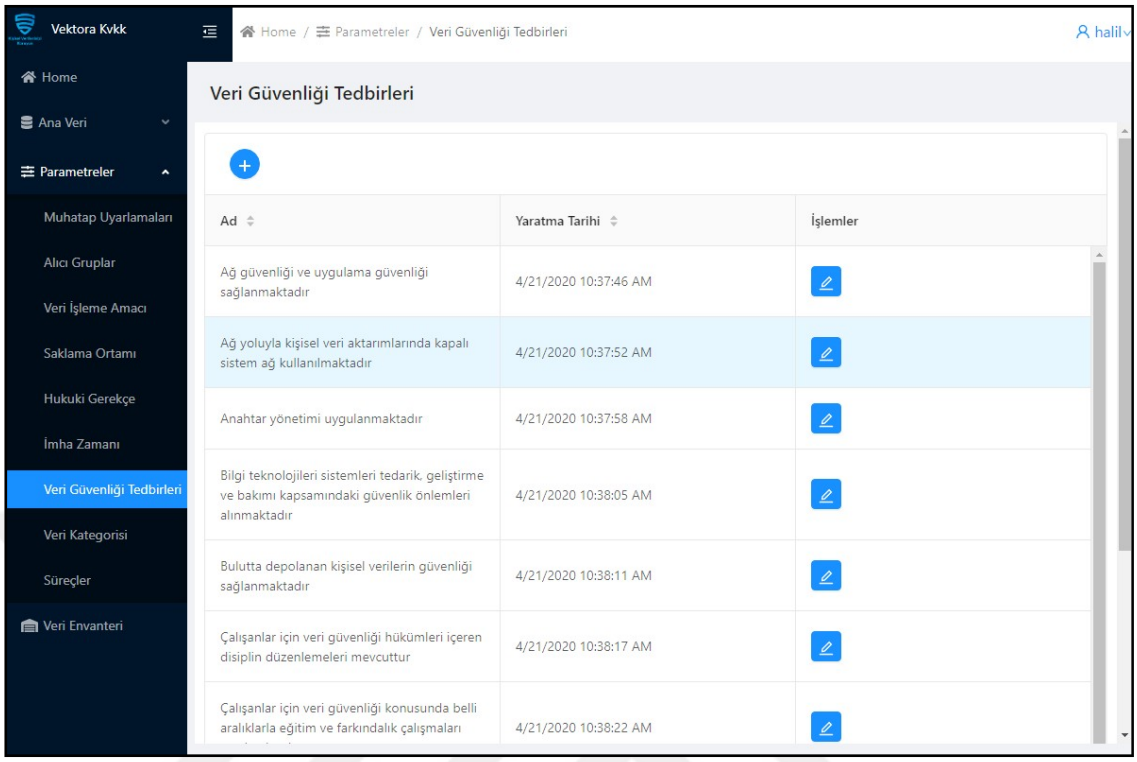
Figure 3.45: Application Destruction Time Page



3.6.12 Data Security Precaution

The security of the personal data is more important than the use of it within the company. Therefore, data security measures taken within the scope of KVKK should be presented to the auditor at the time of the audit. That is why we have to state how we protect the data in the company.

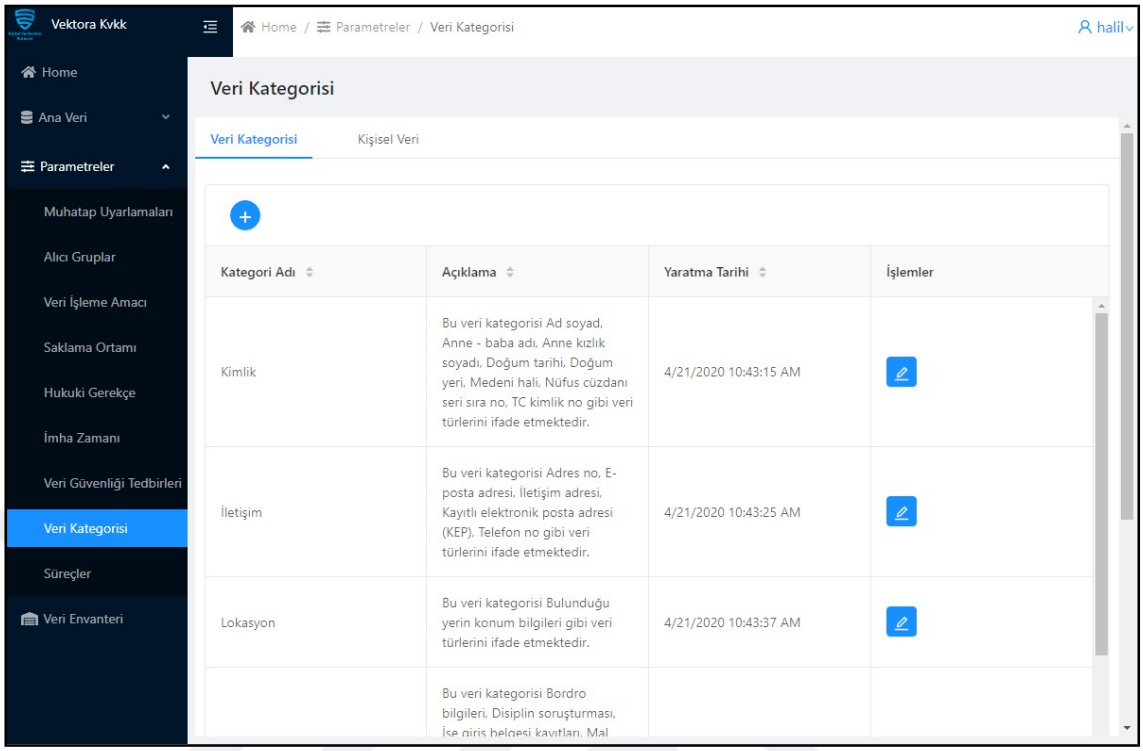
Figure 3.46: Application Data Security Precaution Page



3.6.13 Data Category

In KVKK processes, it is necessary to categorize personal data, so this screen is needed, so the name and full description of the data category should be specified sharply.

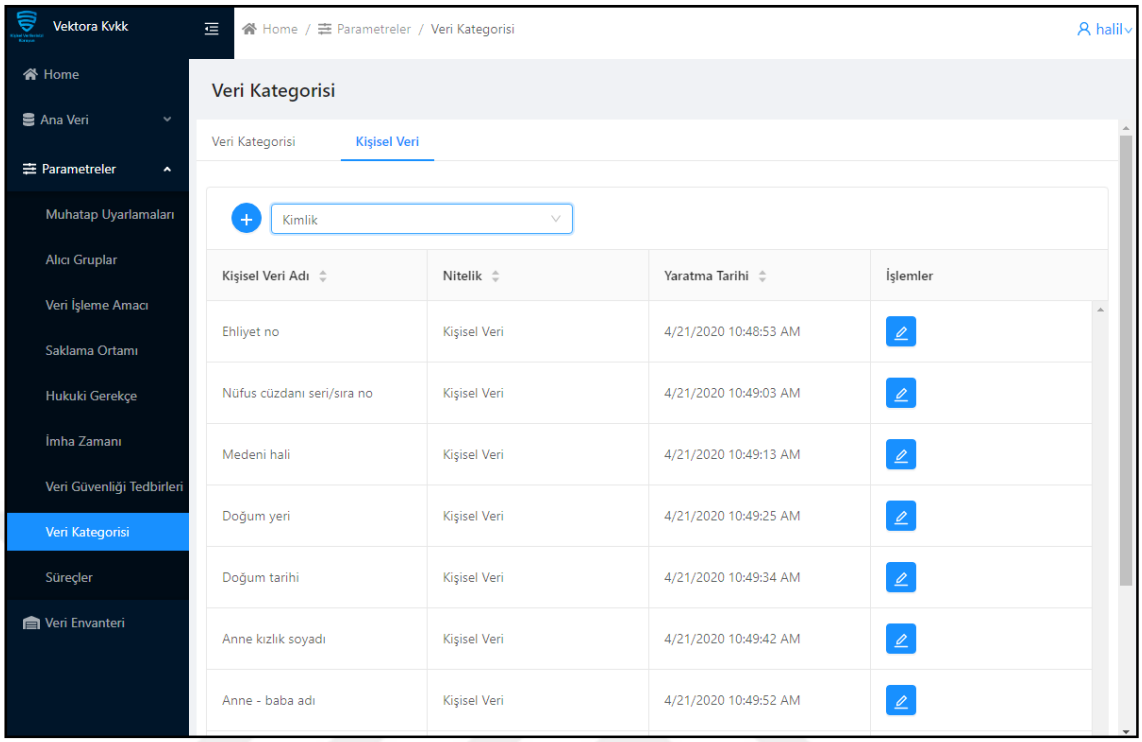
Figure 3.47: Application Data Category Page



3.6.14 Data Category Subset

We categorized personal data, we define personal data in this subset under this category from this screen.

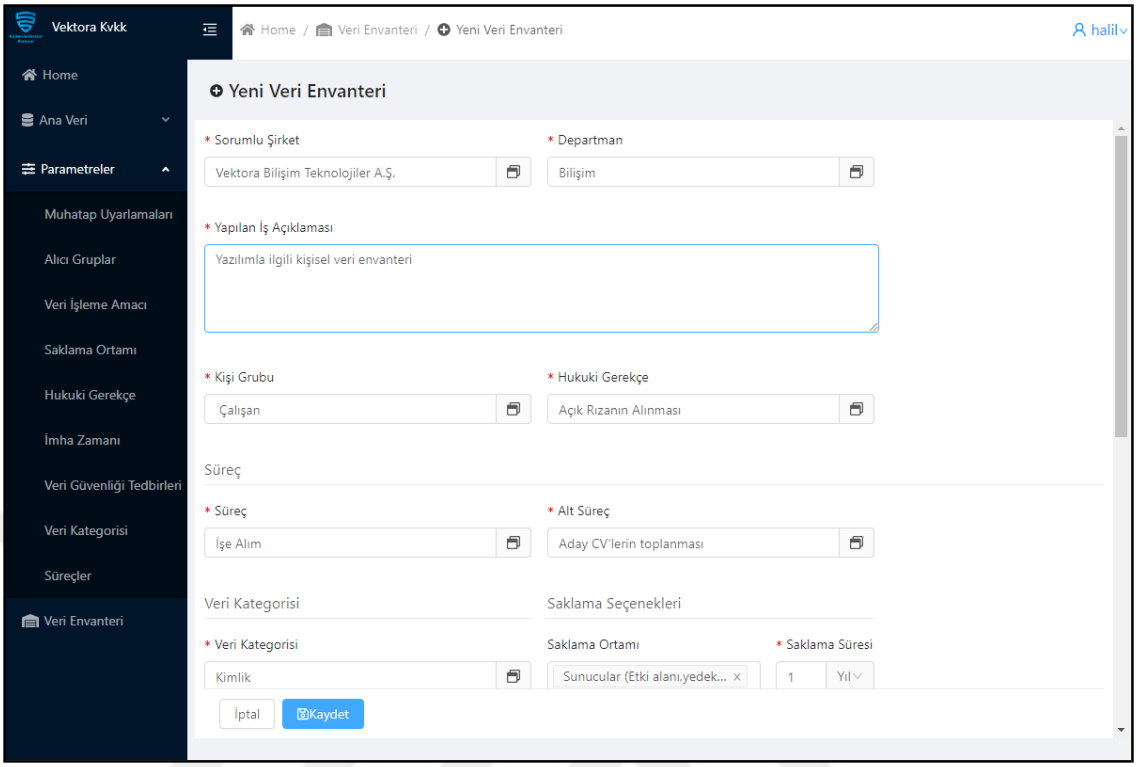
Figure 3.48: Application Data Category Subset Page



3.6.15 Data Inventory

Personal data processing activities carried out depending on the business processes of data officers; personal data processing purposes and legal reason, the data category, the group of recipients transferred and the data subject group, and the personal data they have detailed for the purposes for which the personal data is foreseen, and the data security measures that are foreseen to be transferred to foreign countries. Called data processing inventory.

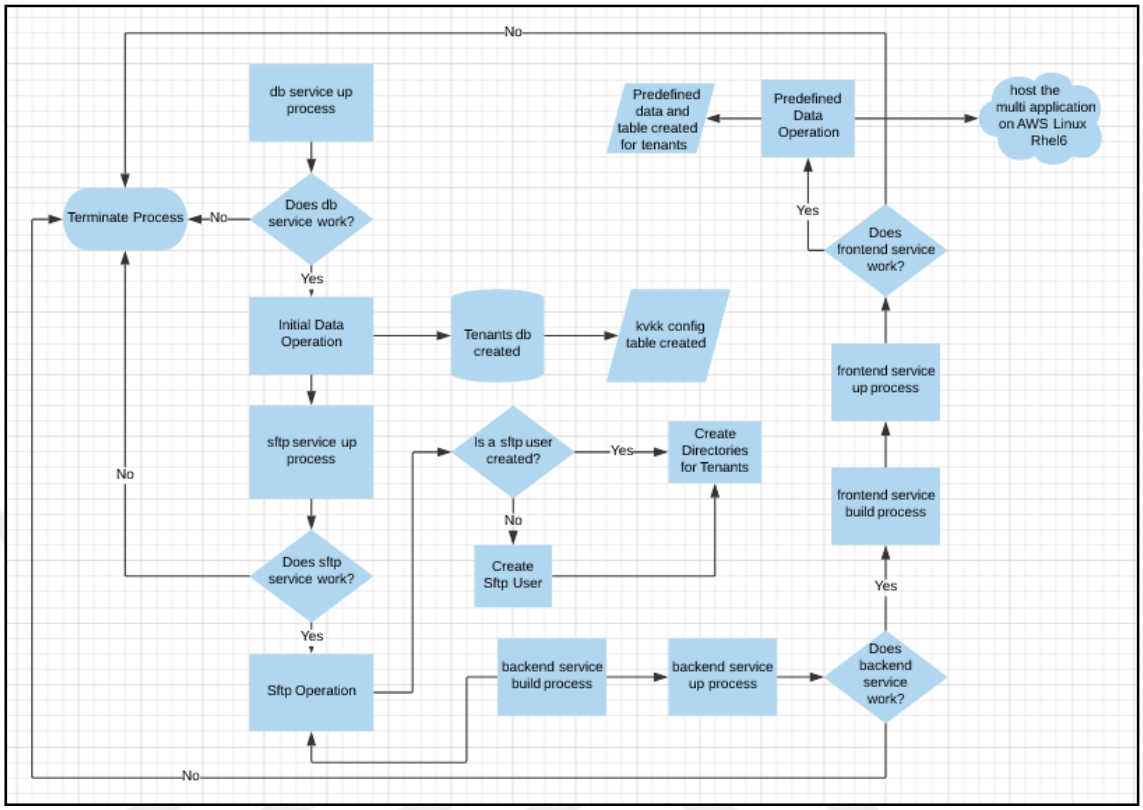
Figure 3.49: Application Data Inventory Page



3.7 KVKK MULTI APPLICATION EXECUTION FLOW CHART

In this section, a flow chart with all the processes of KVKK multiple application is drawn.

Figure 3.50: KVKK Multi Application Execution Process.



3.8 COMPARISON ON BILIMP AND THIS APPLICATION

Bilimp is a corporate web application where KVKK processes are processed. Likewise, it supports multi tenancy structure. Since the back side of the application cannot be predicted, there will be a comparison only on the front side.

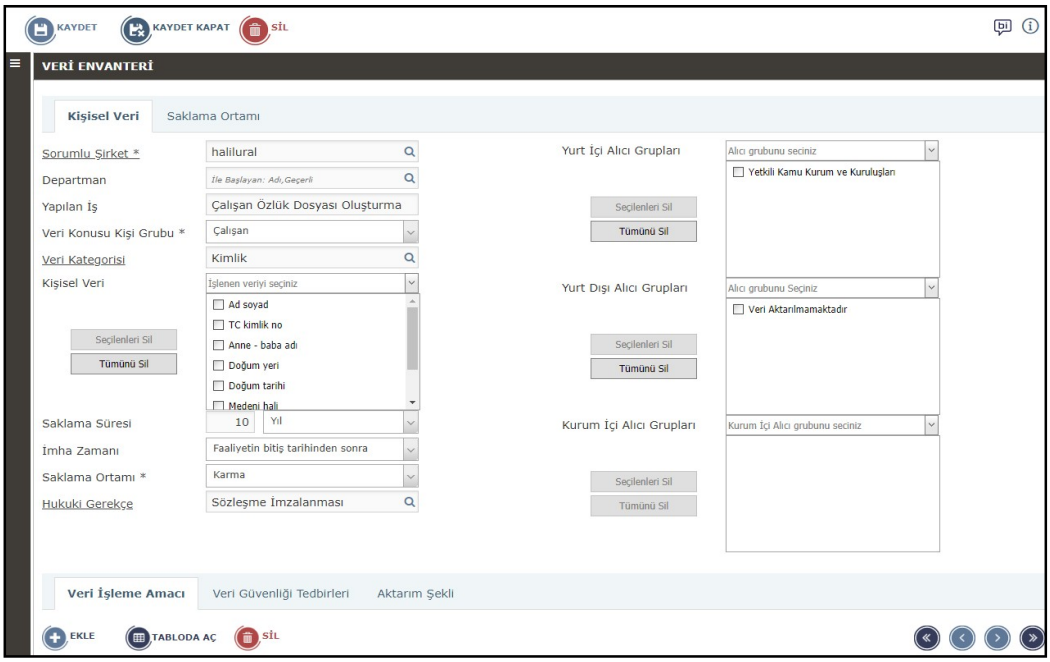
Table 3.7: Comparison on Bilimp and This Application

	Bilimp	**This Application**
Frontend Framework	ASP.net Web Forms	React
Backend Framework	ASP.net	JAVA Spring Framework
Hosting	Only Windows Server	Host Independent
Dockerization	No	Yes
KVKK Wizard	Yes	No
Routing	Different Page	Same Page
HR Process	Yes	Just KVKK
Responsiveness	No	Yes

In the following two applications, images are shared from only one screen, both the list page and the data entry page are available.

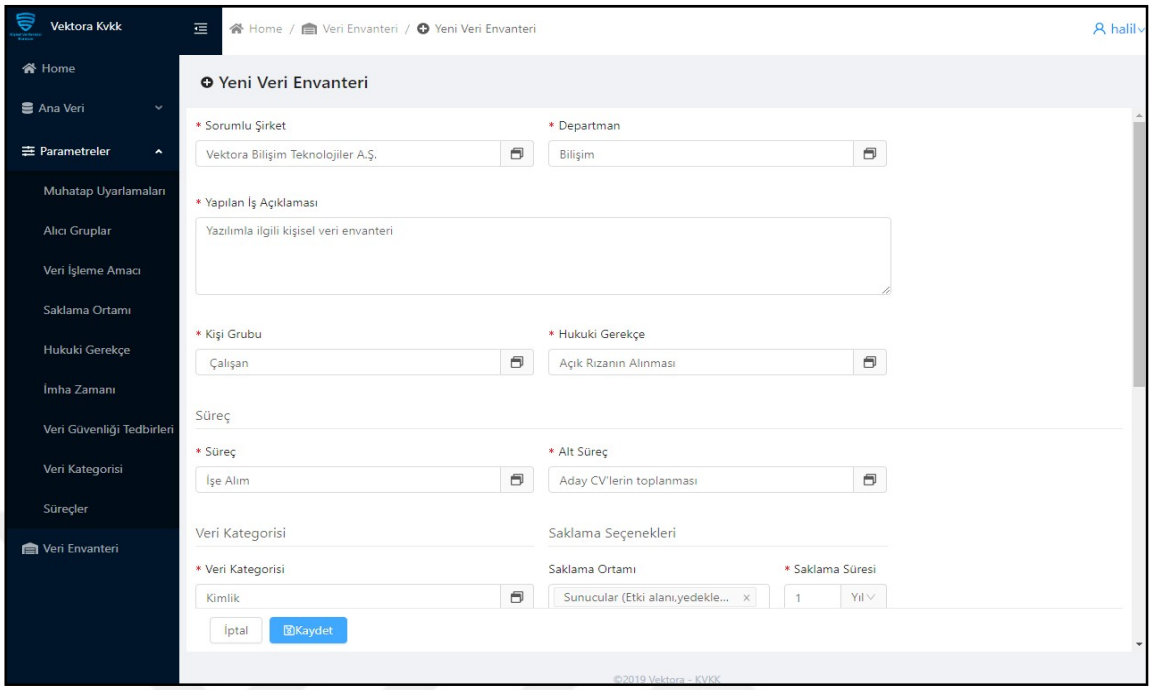
Bilimp application looks like below,

Figure 3.51: Bilimp Application Data Inventory Page



This application looks like below.

Figure 3.52: KVKK Application Data Inventory Page



4. FINDINGS

4.1. SURVEY METHOD AND EVALUATION OF DATA

KVKK adequacy on the website of the participants in the study was aimed to determine the level of overall usability of the site. In addition, it will be examined whether the evaluations about the competence of the website on KVKK and the general usability levels of the website differ according to the demographic characteristics of the participants. It was investigated whether the website's adequacy in KVKK and the general usability levels of the website are affected by factors such as gender, marital status, occupation, education, occupation, Web knowledge and KVKK knowledge.

The hypotheses developed in this context are as follows.

H1: Participants' genders affect the website's proficiency assessments on KVKK.

H2: The sex of the participants affects the KVKK usability assessments of the website.

H3: The marital status of the participants affects the proficiency assessments of the website on KVKK.

H4: The marital status of the participants affects the KVKK usability evaluations of the website.

H5: The educational status of the participants affects the proficiency assessments of the website on KVKK.

H6: The educational status of the participants affects the KVKK usability evaluations of the website.

H7: Participants' knowledge of the website design and their knowledge affect the proficiency assessments of the website about KVKK.

H8: Participants' knowledge of website design and knowledge affect the website's KVKK usability evaluations.

H9: The knowledge of the participants about KVKK affects the proficiency evaluations of the website regarding KVKK.

H10: Participants' knowledge of KVKK affects the KVKK usability evaluations of the website.

H11: Participants' professional status affects the proficiency assessments of the website regarding KVKK.

H12: Participants' professional status affects the KVKK usability evaluations of the website.

H13: Participants' monthly income status affects the proficiency assessments of the website regarding KVKK.

H14: Participants' monthly income status affects the KVKK usability evaluations of the website.

the investigation of opinions about the website of the participants in the study survey method was used. The site's usability level was developed by Brooke (1996) and Kadiran (2015) used a scale of 10 expressions in which Turkish reliability studies were performed. Since there is no ready scale regarding the competence in KVKK and the competency levels of the site, the expressions were formed from the subject headings.

The application was applied to the participants who spent at least half an hour on the website between April and May 2020 and who examined the site completely. The application was made with the online questionnaire system and the data were recorded.

Descriptive statistics of the participants were given with mean, standard deviation, frequency and percentage values. Scales were examined by reliability analysis and factor analysis, and dimensions were determined. T-test analysis and Anova test were performed to examine the evaluation and qualification scores according to the demographic and other characteristics of the participants. Correlation analysis was applied to determine the relationship between the dimensions. Critical decision making level is accepted as 0,05 in the analysis and the analyzes are calculated with SPSS 22.00 program.

Table 4.1: Characteristics Of The Participants

Katılımcıların Özellikleri		n	%
Gender	Male	68	51,5%
	Female	64	48,5%
Age	18-25	16	12,1%
	26-35	78	59,1%
	36-45	18	13,6%
	46 yaş ve üzeri	20	15,2%
Marital status	married	68	51,5%
	Singel	64	48,5%
Education	University	112	84,8%
	Graduate	20	15,2%
İncome	Under 5000TL	48	36,4%
	5001-7500 TL	38	28,8%
	Above 7501 TL	46	34,8%
Technical information about Website Design	Yes	70	53,0%
	No	62	47,0%
Computer, Software engineer and other professional profession in this field	Yes	70	53,0%
	No	62	47,0%
Adequate information on personal data protection law	Yes	76	57,6%
	No	56	42,4%

In the study, 52% of the participants were male and 48% were female. It was observed that 12% of the participants were 18-25 years old, 59% were 26-35 years old, 14% were 36-45 years old and 15% were 46 years old and above. It was determined that 52% of the participants were married and 48% were single. It was determined that 85% of the participants had undergraduate and 15% graduate education.

It was observed that 53% of the participants had a technical level of knowledge about the WEB site design. It was observed that 53% of the participants had a Computer, Software engineer and other professional professions in this field. 58% of the participants were found to have sufficient knowledge about the Law on the protection of Personal data.

4.2 EXAMINING THE SCALES

4.2.1 Examining The Usability Statements Of The Site

Work on the 10-item scale used to investigate the level of website usability has been shown to have a high level of overall confidence level. It was observed that the reliability coefficient of the scale was at the level of 0.90 and the expressions in the scale were gathered under a single dimension. It was determined that the explained variance level was 68% and the KMO sample adequacy level was 0.89. According to these results, the scale was found to be generally consistent measurement tool.

Table 4.2: Examining the Usability Statements of the Site

İfadeler	X	s.s.
1. I think I will use this system frequently.	3,91	1,00
2. I found the system unnecessarily complicated.	3,17	1,04
3. I thought the system was easy to use.	4,09	0,90
4. I think I may need the support of a technical person to use this system.	3,94	1,29
5. I found the various functions in the system well integrated.	3,95	1,14
6. I thought there was too much inconsistency in the system.	3,33	1,16
7. I think many people can use this system quickly.	3,97	1,06

8. I found the use of the system too cumbersome.	3,32	1,17
9. I was sure of myself while using the system.	3,91	1,03
10. I had to learn many things before logging on to the system.	3,83	1,31

In general the assessment made of the respondents said that in the medium and high levels. When the expressions were examined, the highest usability scores of the participants were "I thought the system was easy to use." It was observed that he was in his statements. The lowest ratings were "I thought there was too much inconsistency in the system.", "I found the system unnecessarily complicated." It was observed that he was in his statements. It has been observed that all of the overall scores are not low and the usability levels of the site are medium level and above.

4.2.2 Examining The Statements Of The Site Regarding The Adequacy Of The Kvkk Process

In the study, it was observed that the 14-item scale used to examine the adequacy levels of the website in the KVKK process generally has a high level of reliability. The reliability coefficient of the scale was found to be 0.92 and the expressions in the scale were gathered under a single dimension. It was determined that the explained variance level was 72% and the KMO sample adequacy level was 0.91. According to these results, the scale was found to be a generally consistent measurement tool.

Table 4.3: Examining the Statements of the Site regarding the adequacy of the KVKK process

expressions	X	s.s.
1. The company is sufficient in terms of introducing the main data such as Department, Position, and Persons.	4,09	0,98
2. It is sufficient to make adaptive adaptations.	4,06	1,03
3. Creating recipient groups is easy and convenient	4,15	1,01
4. It is sufficient to enter Data Processing Purposes.	4,11	0,96
5. Sufficient for the creation of Storage Environments. "	4,12	0,98

6. It is sufficient in terms of writing legal reasons.	4,11	0,99
7. It is sufficient to write personal data destruction times	4,03	1,02
8. It is sufficient to enter the Data Security Importance.	4,12	0,97
9. The site is sufficient for entering personal data and personal data subsets.	4,00	1,06
10. The site is sufficient in terms of creating data inventories.	4,15	0,98
11. Sign up page. Is sufficient for a password policy that can prevent weak passwords.	4,02	1,06
12. Being Sufficient in Updating Personal Data and Providing Accuracy	4,05	1,05
13. Sufficiency of Personal Data in that it is Limited, Associated and Metered for the Purpose for Processing	4,06	1,00
14. It is sufficient to process Personal Data in accordance with Law and Integrity Rules.	4,03	0,99

Evaluation in all of the statements are made fairly high scores, the statement said that the high levels of competence in the process of some KVKK site. Statements with the highest scoring of the participants are “Easy and convenient to create recipient groups”, “The site is sufficient in terms of creating data inventories”, “It is sufficient in terms of entering Data Security Importance”, “It is sufficient in terms of writing legal grounds.” There were expressions. Although the rating scores of the statements “The site is sufficient in terms of entering personal data and personal data subsets” and “sign up page is sufficient for a password policy that can prevent weak passwords” are generally high, they have the lowest score among the statements.

Table 4.4: Examination of the expressions related to the general sufficiency of the site

expressions	**X**	**s.s.**
1. It is sufficient for the application to be responsive (compatibility with mobile, tablet and PC screens).	3,89	1,09

2. Generally good in terms of Web design.	4,06	1,03
3. The accessibility of the site is good.	4,06	1,05
4. Problems or unwanted errors appear as usage on the site.	3,26	1,43
5. I am of the opinion of a suitable website in terms of managing KVKKK processes	4,05	1,08
6. The site is sufficient in terms of security (https on the left, can be given as an example, meaning secure.	4,06	1,12

In the study, it was observed that the 6-item scale used to examine the overall competence levels of the website has a high level of reliability in general. The reliability coefficient of the scale was found to be 0.81 and the expressions in the scale were gathered under a single dimension. It was determined that the explained variance level was 65% and the KMO sample adequacy level was 0.80. According to these results, the scale was found to be a generally consistent measurement tool.

In general, it can be said that the evaluation scores are medium and high. “It is sufficient for the application to be responsive (compatibility with mobile, tablet and PC screens).” And "problems or unwanted errors occur as usage on the site." Their expressions are seen as the ones with the lowest scores. Other expressions are at similar levels and appear to have high scores.

Table 4.5: Examining dimension scores in general

Scale	X	s.s.
Site Usability	3,24	0,61
Adequacy of the site in terms of KVKK processes	4,08	0,91
General Sufficiency of the Site	3,89	0,93

In the study, it was determined that participants' Site Usability levels were 3.24 ± 0.61, and the site's adequacy in terms of KVKK processes was 4.08 ± 0.91. The overall

sufficiency of the site was determined to be 3.89 ± 0.93. Generally, it can be said that the usability level of the site is at medium levels, and in terms of KVKK processes and its general level is sufficient.

4.3 EXAMINATION OF USABILITY AND COMPETENCE ACCORDING TO THE CHARACTERISTICS OF THE PARTICIPANTS

Table 4.6: Examining Usability and Sufficiency by Gender

Gender Dimensions	Gender Dimensions	n	X	s.s.	p
Site Usability Male Woman	Male	68	3,34	0,50	0,04*
	Woman	64	3,14	0,70	
Adequacy of the site in terms of KVKK processes Male Woman	Male	68	4,30	0,81	0,01*
	Woman	64	3,84	0,97	
General Sufficiency of the Site Male	Male	68	4,08	0,87	0,01*
	Woman	64	3,69	0,94	

It was seen that the usability levels of the participants were at different levels according to their gender. It was found that the reason for the difference was that the usability scores of male participants were higher than women. It was found that male participants think that the site can be used at higher levels than women (p = 0.04).

It was seen that the level of proficiency of the participants in terms of protecting the personal data of the site is different according to their gender. It was determined that the difference was due to the higher level of competence scores of male participants in terms of protecting the personal data of the site compared to women. Male participants think that the site is more efficient in terms of protecting personal data than women (p = 0.01).

It was seen that the general competence levels of the participants were at different levels according to their gender. It was determined that the difference was due to the fact that

the overall competence scores of male participants were higher than women. Male participants think that their general competence is higher than women (p = 0.01).

Table 4.7: Examining Usability and Sufficiency by Marital Status

Dimensions Marital Status	Dimensions Marital Status	n	X	s.s.	p
Site Usability Married Single	Married	68	3,22	0,53	0,63
	Single	64	3,27	0,69	
Sufficiency of the site in terms of KVKK processes Married Single	Married	68	4,05	0,83	0,70
	Single	64	4,11	1,00	
General Competence of the Site Married	Married	68	3,86	0,88	0,64
	Single	64	3,93	0,98	

It was observed that the usability level scores of the participants were not at different levels according to the marital status, and the usability assessment levels of the married and single participants were similar (p = 0.63).

It was observed that the proficiency level scores of the participants in terms of KVKK processes were not at different levels according to the marital status, and that the married and single participants' proficiency assessment levels in terms of the KVKK processes were similar (p = 0.70).

It was observed that the general competence level scores of the participants were not at different levels according to the marital status, and the general competence evaluation levels of the married and single participants were similar (p = 0.64).

Table 4.8: Examining Usability and Sufficiency by Marital Status

scale	scale Education	n	X	s.s.	p
Site Usability License Graduate	License	112	3,16	0,54	0,01*
	Graduate	20	3,71	0,77	

Adequacy of the site in terms of KVKK processes License Graduate	License	112	4,05	0,96	0,04*
	Graduate	20	4,24	0,54	
General Sufficiency of the Site License	License	112	3,82	0,96	0,03*
	Graduate	20	4,32	0,56	

It has been seen that the usability levels of the participants differ according to their education levels. It was determined that the difference was due to the fact that the usability points of the participants having postgraduate education were higher than the graduates. Participants with postgraduate education were found to think that the site could be used at higher levels than undergraduate graduates (p = 0.01).

It has been seen that the competence levels of the participants in terms of protection of personal data are at different levels according to their education levels. It was determined that the difference was due to the fact that the participants who had undergraduate education had higher proficiency scores in terms of protecting the personal data of the site than the individuals who had graduate education. Participants at the undergraduate level think that the site is more efficient in terms of protecting personal data than those who have postgraduate education (p = 0.04).

It was seen that the general competence levels of the participants differed according to their education levels. It was determined that the difference was due to the fact that the general qualification scores of the participants who had graduate education were higher than the graduates. Participants with postgraduate education were found to think that the site was sufficient at higher levels than undergraduate graduates (p = 0.03).

Table 4.9: Investigation of Usability and Sufficiency According to Technical Knowledge of Website Design

Scale	Technical information about Dimensions Website Design	n	X	s.s.	p
Site Usability License	Yes	70	3,46	0,47	0,01*

Graduate	No	62	2,99	0,66	
Adequacy of the site in terms of KVKK processes License Graduate	Yes	70	4,34	0,65	0,01*
	No	62	3,78	1,07	
General Sufficiency of the Site License	Yes	70	4,14	0,73	0,01*
	No	62	3,61	1,04	

It has been seen that the usability levels of the participants differ according to their education levels. It was determined that the difference was due to the fact that the usability points of the participants having postgraduate education were higher than the graduates. Participants with postgraduate education were found to think that the site could be used at higher levels than undergraduate graduates (p = 0.01).

It has been seen that the competence levels of the participants in terms of protection of personal data are at different levels according to their education levels. It was determined that the difference was due to the fact that the participants who had undergraduate education had higher proficiency scores in terms of protecting the personal data of the site than the individuals who had graduate education. The participants at the undergraduate level think that the site is highly qualified in terms of protecting personal data compared to those who have postgraduate education (p = 0.04).

It was seen that the general competence levels of the participants differed according to their education levels. It was determined that the difference was due to the fact that the general qualification scores of the participants who had graduate education were higher than the graduates. Participants with postgraduate education were found to think that the site was sufficient at higher levels than undergraduate graduates (p = 0.03).

Table 4.10: Investigation of Usability and Competency According to Website Design and Professional Profession

Dimensions	**Computer, Software engineer and other professional in this field**	**n**	**X**	**s.s.**	**p**
Site Usability	Yes	70	3,34	0,61	0,04*
	No	62	3,13	0,59	

Adequacy of the site in terms of KVKK processes	Yes	70	4,25	0,80	0,02*
	No	62	3,88	1,00	
General Sufficiency of the Site	Yes	70	4,02	0,84	0,04*
	No	62	3,75	1,00	

It has been seen that the usability levels of the participants differ according to the professional status of the profession. It was determined that the difference was due to the fact that the usability scores of the participants with professional occupations were higher than those without professional occupations. It was determined that the participants with professional professions thought that the site could be used at higher levels than the non-participants (p = 0.01).

It was seen that the competence levels of the participants in terms of protecting the personal data of the site are at different levels according to the professional status of the profession. It has been determined that the participants who have professional professions are higher than the individuals who do not have professional professions in terms of protecting the personal data of the site (p = 0.01).

It was seen that the general competence levels of the participants were at different levels according to the professional professions. It was determined that the difference was due to the fact that the participants with professional professions had a higher level of general competence scores of the site compared to individuals without technical knowledge (p = 0.01).

Table 4.11: Investigation of Usability and Adequacy by Website Design and KVKK Knowledge Level

Dimensions	**s Adequate information on the personal data protection law**	**n**	**X**	**s.s.**	**p**
Site Usability	Yes	76	3,32	0,55	0,10
	No	56	3,14	0,67	

Adequacy of the site in terms of KVKK processes	Yes	76	4,14	0,86	0,36
	No	56	3,99	0,99	
General Sufficiency of the Site	Yes	76	3,96	0,89	0,37
	No	56	3,81	0,97	

It was observed that the participants' knowledge about the site's usability level scores were not at different levels depending on the situation, and the usability evaluation levels of the participants with or without KVKK were similar (p = 0.10).

It was observed that the level of proficiency of the participants in terms of KVKK processes was not at different levels according to their knowledge about KVKK, and the level of proficiency evaluation of the participants with or without KVKK in terms of KVKK processes (p = 0.36).

It was observed that the general competence level scores of the participants were not at different levels according to their knowledge about KVKK, and the general competence evaluation levels of the participants with or without KVKK were similar (p = 0.37).

Table 4.12: Investigation of Usability and Competence by Website Design and Income Level

Dimensions income	**Dimensions income**	**n**	**X**	**s.s.**	**p**
Site Usability	Under 5000TL	48	3,41	0,65	
	5001-7500 TL	38	3,04	0,56	0,05
	Over 7501 TL	46	3,23	0,57	
The site's adequacy in terms of KVKK processes	below 5000TL	48	3,98	0,84	
	5001-7500 TL	38	4,11	1,04	0,63
	Over 7501 TL	46	4,16	0,89	
General Sufficiency of the Site is	below 5000TL	48	3,86	0,88	
	5001-7500 TL	38	3,89	0,98	0,94
	Over 7501 TL	46	3,93	0,94	

It was observed that the usability level scores of the participants were not at different levels according to the income situation, and the usability assessment levels of the participants with monthly income below 5000 TL, between 5000-7500 TL or above 7501 TL (p = 0.05).

It was seen that the level of proficiency of the participants in terms of KVKK processes is not different than the monthly income level, and the level of proficiency evaluation of the site in terms of KVKK processes of the participants with monthly income below 5000 TL, between 5000-7500 TL or over 7501 TL (p = 0.63). .

It was observed that the general proficiency level scores of the participants were not at different levels according to their monthly income status, and the general proficiency evaluation levels of the participants with monthly income below 5000 TL, between 5000-7500 TL or above 7501 TL (p = 0.94).

Table 4.13: Investigation of the Relationship Between Site Usability, KVKK Processes and General Adequacy of the Site

Boyutlar		**Site Usability**	**The site's adequacy in terms of KVKK processes**	**General Sufficiency of the Site**
Site Usability	r	1		
	p			
The site's adequacy in terms of KVKK processes	r	0,60*	1	
	p	0,01		
General Sufficiency of the Site	r	0,63*	0,90*	1
	p	0,01	0,01	

It was observed that there was a positive and strong relationship between the Usability Scores of the participants and the proficiency scores of the Site in terms of KVKK processes (r = 0.60, p = 0.01).

It was observed that there was a positive and strong relationship between the Usability Scores of the participants and the overall competence scores of the Site ($r = 0.63$, $p = 0.01$).

It was observed that the participants had a positive and very strong relationship between the adequacy of the Site in terms of KVKK processes and the overall competence scores of the Site ($r = 0.90$, $p = 0.01$).

In general, the participants who found the site usable were also found to think that the site was sufficient in terms of KVKK processes and sufficient in terms of general features.

5. CONCLUSION AND DISCUSSION

In general, all system are designed and containerized with Docker. The Micro services such as, MySQL, backend, frontend and SFTP runs as a container and isolated from each other. Frontend cannot connect to the MySQL and SFTP services. The structure of the system is available to work with multiple customers. Customers cannot access each other's databases because their databases are completely isolated and operate on the same hardware. If desired, any customer's database can be made operational on their own systems, and there is a margin of flexibility in this respect.

In this study, especially, I realized the systematic working of all applications on a server. Docker is a cut out for this. The problem is the lack of processing, managing and maintaining KVKK processes. Regarding this problem, an advanced corporate web application has been designed and launched. A high degree of importance was attached to security, and the SSL certificate was signed for applications running on the backend and frontend operating on servers. SSL certificates were signed on the tomcat side for the backend and on the NGINX side for the frontend. As the databases are completely isolated from each other, customers will be able to securely process and organize their personal data.

In the same way, on the document management side, if the user wants, she/he can manage her/his documents using her own SFTP servers. The system provides this flexibility to users. Since JWT is used on the backend side, security has been completely taken into account because the top layer is a security measure, because the port that opens out is 8443 and the https certificate is signed. Also, at the point of calling backend API services, role authorization can be made for users, thus preventing unwanted access.

As with all systems, this system also has disadvantages. One of the disadvantages of the system is the multitenant database structure, so the database table and field updates are handled manually. Spring Hibernate framework, which updates the basic database, was used in the backend system. Unfortunately, this framework does not update customers'

database table and field maintenance, this has to be done manually. This creates a separate effort.

Another disadvantage is that there is no structure that logs http 500 errors on the backend side. This is a critical need for monitoring, interpreting and analyzing errors, especially if we consider that different error scenarios will exist in different customers.

Another disadvantage is that a strong backup system, especially of all databases, is not installed. In this sense, if the backup system is installed and there is a problem in the current databases, it is necessary to return from the last backup.

In the study, examining the adequacy and usability of the KVKK website was observed in the study, 52% of the participants were male and 48% were female. It was observed that 12% of the participants were 18-25 years old, 59% were 26-35 years old, 14% were 36-45 years old and 15% were 46 years old and above. It was determined that 52% of the participants were married and 48% were single. It was determined that 85% of the participants had undergraduate and 15% graduate education. It was observed that 53% of the participants had a technical level of knowledge about the WEB site design. It was observed that 53% of the participants had a Computer, Software engineer and other professional professions in this field. 58% of the participants were found to have sufficient knowledge about the Law on the protection of Personal data.

In the evaluations made by the participants, it can be said that it is generally at medium and high levels. When the expressions were examined, the highest usability scores of the participants were “I thought the system was easy to use.” It was observed that he was in his statements. The lowest ratings were "I thought there was too much inconsistency in the system.", "I found the system unnecessarily complicated.", "I found the system too cumbersome to use." It was observed that he was in his statements. It has been observed that all of the overall scores are not low and the usability levels of the site are medium level and above.

It can be said that the evaluations of all the statements were made with very high scores, and the competence of the site in the KVKKK process on the basis of expression is at high levels. Statements with the highest scoring of the participants are “Easy and

convenient to create recipient groups", "The site is sufficient for creating data inventories", "Sufficient for entering Data Security Importance", "Sufficient for writing legal grounds." There were expressions. Although the rating scores of the statements "The site is sufficient in terms of entering personal data and personal data subsets" and "sign up page.

In the study, it was determined that participants' Site Usability levels were 3.24 ± 0.61, and the site's adequacy in terms of KVKK processes was 4.08 ± 0.91. The overall sufficiency of the site was determined to be 3.89 ± 0.93. Generally, it can be said that the usability level of the site is at medium levels, and in terms of KVKK processes and its general level of sufficiency is high.

Table 5.1: Evaluation of Hypotheses

Status	Acceptance	Rejection
H1: Participants' genders affect the website's proficiency assessments on KVKK.		x
H2: The sex of the participants affects the KVKK usability assessments of the website.		x
H3: The marital status of the participants affects the proficiency assessments of the website on KVKK.	x	
H4: The marital status of the participants affects the KVKK usability evaluations of the website.	x	
H5: The educational status of the participants affects the proficiency assessments of the website on KVKK.		x
H6: The educational status of the participants affects the KVKK usability evaluations of the website.		x
H7: Participants' knowledge of the website design and their knowledge affect the proficiency assessments of the website about KVKK.		x
H8: Participants' knowledge of website design and knowledge affect the website's KVKK usability evaluations.		x
H9: The knowledge of the participants about KVKK affects the proficiency evaluations of the website regarding KVKK.	x	
H10: Participants' knowledge of KVKK affects the KVKK usability evaluations of the website.	x	
H11: Participants' professional status affects the proficiency assessments of the website regarding KVKK.		x
H12: Participants' professional status affects the KVKK usability evaluations of the website.		x
H13: Participants' monthly income status affects the proficiency assessments of the website regarding KVKK.	x	

H14: Participants' monthly income status affects the KVKK usability evaluations of the website.	x

It was found that male participants think that the site can be used at higher levels than women. Male participants think that the site is high enough to protect personal data compared to women. Male participants think that KVKK and general competence are higher than women. H1-H2 is accepted.

It was observed that the evaluation levels of usability, KVKK and general competence of married and single participants were at similar levels. H3-4 is rejected.

It has been seen that the levels of usability, KVKK and general competence evaluation levels of the participants with higher education levels are higher than the graduates. H5-H6 is accepted.

It was observed that the levels of evaluation of usability, KVKK and general competence of the participants who have knowledge about website design were higher than those without knowledge about website design. H7-H8 is accepted.

H9-10 is rejected when their knowledge about KVKK does not have a significant effect on the levels of usability, KVKK and general competence.

It has been observed that the level of evaluation of usability, KVKK and general competence of participants with professional professions is higher than those without professional professions. H11-H12 is accepted.

It was observed that the levels of assessment of usability, KVKK and general competence of the participants with monthly income levels below 5000 TL, 5001-7500 TL and 7501 TL and above were similar. H13-14 is rejected.

6. FUTURE WORK

Currently, only one study for SAP ERP UI-masking is ongoing. Especially, different software systems need to be improved to store, report, monitor, control and anonymize personal data.

RHEL 7 is used on the server side. In personal data security processes, the security vulnerabilities of the server should be fixed, therefore, security vulnerabilities of Lynis and Linux Server, which are security control tools, should be detected and eliminated.

For Docker, which is used to create micro services, Docker Security Benchmark must be realized and vulnerabilities must be resolved on the docker side.

In order to determine the transparency of all KVKK processes in a company and where there are missing points, the KVKK compliance optimization wizard needs to be developed. With this wizard, a company will be able to learn about all KVKK processes end-to-end, analyze where it is missing and take action accordingly, and will be freed from penal sanctions according to KVKK law.

REFERENCES

Books

ACM SIGMOD., 2013. *Management of data*. Vancouver, Canada New York, NY, USA: ACM, pp.1195-1206.

Boettiger, C. (2015). An introduction to Docker for reproducible research. *ACM SIGOPS Operating Systems Review*, 49(1), 71-79.

Bui, T. (2015). Analysis of docker security. ArXiv preprint arXiv:1501.02967.

Chen, Z., 2018. HTML5 Hybrid Mobile Application: Building mobile applications using web technologies with Ionic.

Eisenman, B., 2015. Learning React Native: Building Native Mobile Apps with JavaScript. “O’Reilly Media, Inc."

El Omari, M., Erramdani, M. and Filali, S., 2018 Getting Model of MVVM pattern from UML Models.

Felter, W., Ferreira, A., Rajamony, R., & Rubio, J. (2014). An updated performance comparison of virtual machines and linux containers. Technology, 28, 32.

Gackenheimer C., 2015. Introducing Flux: An Application Architecture for React. In: Introduction to React. Apress, Berkeley, CA

Harji, A. S., Buhr, P. A., & Brecht, T. (2013). Our troubles with Linux Kernel upgrades and why you should care. ACM SIGOPS Operating Systems Review, 47(2), 66-72.

Nikulchev, E., Kolyasnikov, P., Ilin, D., Kasatonov, S., Biryukov, D. and Zakharov, I., 2018, July. Selection of Architectural Concept and Development Technologies for the Implementation of a Web-Based Platform for Psychology

Pastushenko, O., 2017.UXgraph-Vue. Js library with predefined D3 graphs.

Scheepers, M. J. (2014). Virtualization and containerization of application infrastructure: A comparison.

Seo, K.-T., Hwang, H.-S., Moon, I.-Y., Kwon, O.-Y., & Kim, B.-J. (2014). Performance Comparison Analysis of Linux Container and Virtual Machine for Building Cloud.

Turnbull, J. (2014). The Docker Book: Containerization is the new virtualization.

Van der Aalst, W., Weijters, T., & Maruster, L. (2004). Workflow mining: Discovering process models from event logs. Knowledge and Data Engineering, IEEE Transactions on, 16(9), 1128-1142.

Vase, T. (2015). Advantages of Docker.

Waldspurger, C. A. (2002). Memory resource management in VMware ESX server. ACM SIGOPS Operating Systems Review, 36(SI), 181-194.

Clifton C., Housman E., Rosenthal A. (1997), Experience with a combined approach to attribute-matching across heterogeneous databases, in 7th IFIP 2.6 Working Conference on Database Semantics. Leysin, Switzerland: Chapman & Hall, pp. 428–451.

Vaidya J. Clifton C. 2004, Privacy preserving na¨ıve bayes classifier for vertically partitioned data, in 2004 SIAM International Conference on Data Mining, Lake Buena Vista, Florida.

Elfeky M., Verykios V., Elmagarmid A. 2002, TAILOR: A record linkage toolbox, in Proceedings of the 18th International Conference on Data Engineering, San Jose, California.

Belenkiy M., Camenisch J., Chase M., Kohlweiss M., Lysyanskaya A., and Shacham H., 2009, Randomizable proofs and delegatable anonymous credentials. In Proc. Of the 29th Annual International Cryptology Conference on Advances in Cryptology (CRYPTO'09), Santa Barbara, California, USA, LNCS, volume 5677, pages 108–125. Springer-Verlag.

Olumofin F. and Goldberg I., 2010, Privacy-preserving queries over relational databases. In Proc. of the 10th International Conference on Privacy Enhancing Technologies (PETS'10), Berlin, Germany, LNCS, volume 6205, pages 75–92. Springer-Verlag.

Ostrovsky R., Sahai A., and Waters B., (2007), Attribute-based encryption with non-monotonic access structures. In Proc. of the 14th ACM Conference on Computer and Communications Security (CCS'07), Alexandria, Virginia, USA, pages 195–203. ACM.

Paillier P., 1999 Public-key cryptosystems based on composite degree residuosity classes. In Proc. of the 18th International Conference on Theory and Application of Cryptographic Techniques (EUROCRYPT'99), Prague, Czech Republic, LNCS, volume 1592, pages 223–238. Springer-Verlag.

Periodicals

Dawoud, M., (2017) Privacy Preserving Search and Data Retrieval from Clouds. *Thesis for PhD Degree.* Istanbul Technical University FBE.

Koç, H.K., (2019). Biyometrik tanı yöntemlerinde kişisel very güvenliği arttırılmış sistem tasarımı. *Thesis for the Master Degree.* Mersin University FBE.

Mobayenjarihani M., (2018) Quantifying and Protecting Genomic Privacy. *Thesis for the Master Degree.* Bilkent University FBE.

Ramos, M., Valente, M.T. and Terra, R., 2018. AngularJS performance: A survey study. IEEE Software, 35(2), pp.72-79.

Yılmaz, B., (.2019). Türk Anayasa Mahkemesi ve Avrupa İnsan Hakları Mahkemesi ışığında Kişisel Verilerin Korunması. *Thesis for PhD. Degree.* Ankara: Hacettepe SBE.

Other Sources

Angular Official Blog., 2018, July. AngularJS to Angular Concepts: Quick Reference From https://angular.io/guide/ajs-quick-reference

Apache Tomcat, 2020 *SSL/TLS Configuration HOW-TO*. [internet]

Aulbach S., Grust T., Jacobs D., Kemper A. and Rittinger J. 2008, [online] Multi-tenant databases for software as a service: schema-mapping techniques Proceedings of the 2008

Banville R., Holzel R. 2012, [online] Openedge multi-tenancy overview, Progress software. Bedford, http://www.progress.com [accessed 25 April 2020].

DockerHub. 2019, [online] Nginx docker images. https://hub.docker.com. [accessed 23 May, 2019].

Elmore A. J., Das S., Agrawal D. and Abbadi A. 2011 [online] Towards an elastic and autonomic multitenant database Proc. of NetDB Workshop.

European Convention on Human Rights (1950 s. k). **Secretary General of the Council of Europe**, December 11, 1950 Rome.

Grund M., Schapranow M., Krueger J., Schaffner J. and Bog A. 2008, [online] Shared Table Access Pattern Analysis for Multi-Tenant Applications Advanced Management of Information for Globalized Enterprises, 2008. AMIGE 2008. IEEE Symposium on. Pp.1-5.

http://nmhukukburosu.com/general-data-protection-regulation-gdpr-genel-veri-koruma-tuzugu-gvkt-yururluge-girdi/ [accessed 24/04/2020].

https://tomcat.apache.org/tomcat-7.0-doc/ssl-howto.html [accessed 23 April 2020].

Joy, A. M. (2015). Performance comparison between Linux containers and virtual machines. Paper presented at the Computer Engineering and Applications (ICACEA), 2015 International Conference on Advances in.

Kişisel Verileri Koruma Kanunu, 2016.

KVKK, *Ulusal ve Uluslararası Alanda Kişisel Verilerin Korunmasına Duyulan İhtiyaç*, 2018. https://kvkk.gov.tr/ [access date 23/04/2020]

Moreton R., Dudley C., Olumuyiwa M. 2014. A review of multi-tenant Database and Factors that Influence its Adoption. *ResearchGate.*

Schiller O., Schiller B., Brodt A. and Mitschang, B. 2011 [online]. Native support of multitenancy in RDBMS for software as a service Proceedings of the 14th International Conference on Extending Database Technology. Uppsala, Sweden New York, NY, USA: ACM, pp.117-128.

Stefankrause Developer Team., 2018, December. Results for js web frameworks benchmark. From https://www.stefankrause.net/js-frameworksbenchmark6/webdriver-ts-results/table.html

ÜNAL NİZAMOĞLU, İ., 2018, General Data Protection Regulation, [online] 30 may 2018,

Universal Declaration of Human Rights (1948 s. k). **United Nations General Assembly**, 217 A (III); December 10, 1948

Vue Official Blog., 2018, July. Reactive Data Binding Research. In Science and Information Conference From https://v1.vuejs.org/guide/overview.html

Walraven S., Van Landuyt D., Truyen E., Handekyn K. and Joosen W. 2014, [online] Efficient customization of multi-tenant Software-as-a-Service applications with service lines. Journal of Systems and Software, http://www.sciencedirect.com/science [accessed 28 April 2020].

Adam N. R., Wortmann J. C. 1989, [online] Security-control methods for statistical databases: A comparative study, ACM Computing Surveys, vol. 21, no. 4, http://doi.acm.org/10.1145/76894.76895 [accessed 24 May 2020].

Duncan G. T., Keller-McNulty S. A., Stokes S. L. 2001, [online] Disclosure risk vs. data utility: The r-u confidentiality map, National Institute of Statistical http://www.niss.org/technicalreports/tr121.pdf [accessed 24 May 2020].

Agrawal R., Evfimievski A., Srikant R. 2003, [online] Information sharing across private databases, in Proceedings of ACM SIGMOD International Conference on Management of Data, San Diego, California, June 9-12 http://doi.acm.org/10.1145/872757.872771 [accessed 25 May 2020].

Agrawal R., Kiernan J., Srikant R., Y. Xu 2002 [online], Hippocratic databases, in Proceedings of the 28th International Conference on Very Large Databases, http://www.vldb.org/conf/2002/S05P02.pdf [accessed 25 May 2020].

Cranor L., Langheinrich M., Marchiori M., Presler-Marshall M., Reagle J. 2002, [online] The platform for privacy preferences 1.0 (P3P1.0) specification, http://www.w3.org/TR/P3P/ [accessed 23 May 2020].

Atallah M. J., Elmongui H. G., Deshpande V., Schwarz L. B. 2003, [online] Secure supply-chain protocols," in IEEE International Conference on E-Commerce, http://ieeexplore.ieee.org/xpl/citationdwnld.jsp?arNumber=1210264 [accessed 22 May 2020].

Li W.-S., Clifton C. 2000, [online] SEMINT: A tool for identifying attribute correspondences in heterogeneous databases using neural networks, http://dx.doi.org/10.1016/S0169-023X(99)00044-0 [accessed 22 May 2020].

Doan A., Domingos P., Halevy A. 2003, [online] Learning to match the schemas of databases: A multistrategy approach, Machine Learning Journal, vol. 50, pp. 279–301, http://anhai.cs.uiuc.edu/home/papers/lsd-mlj03.pdf [accessed 22 May 2020].

Castano S. Antonellis V. D. 1999, [online] A schema analysis and reconciliation tool environment, in Proceedings of the Int. Database Engineering and Applications Symposium (IDEAS).

Bethencourt J., Sahai A., and Waters B., 2007, Ciphertext-policy attribute-based encryption. In Proc. Of the 2007 IEEE Symposium on Security and Privacy (S&P'07), Berkeley, California, USA, pages 321–334. IEEE.