Karolinska's IT-landscape

Innehåll

1 Definitions

Term	Definition
Digital Health Platform	Karolinska's new platform that offers integration and data storage services. Consists of a clinical data repository (CDR) based on an international standard (openEHR), development tools and a traditional enterprise data warehouse (EDW). Is used as a base for the realization of applications in a health care setting.
OCP	OpenShift Container Platform (OCP) gives organizations the ability to build, deploy, and scale applications faster.
SF	Region Stockholm's Service Management Department (SF) for common IT-services, such as IT-infrastructure.

2 Background and Context

This document describes Karolinska's strategy for IT and IT-development. The strategy encompasses capabilities, expertise, and infrastructure, with the aim of meeting Karolinska's needs in an agile, secure, and cost-effective manner. The document briefly outlines the target state, strategy, status, and planned transition steps. The transition steps involve increasing levels of detail as uncertainty in planning decreases.

The purpose of this document is to assist decision-makers, suppliers, and others in better understanding where Karolinska's health care organization is headed and how we plan to implement various types of IT functions.

3 Goals and Strategy

To meet regulatory and other legal requirements, IT solutions must be designed based on the type of information they manage. At the same time, Karolinska aims to, quickly and cost-effectively, implement working solutions. To manage both perspectives, we have established a strategy based on an information-centric, standardized technology platform. In this context, we define the concept of *platform* as a common foundation from which Karolinska can rapidly and securely implement solutions.

A standardized technology platform shifts fundamental parts of the design and operation of applications to services within a unified platform. Developers and suppliers can then focus on business needs and, by integrating with and/or consuming services from the platform, realize solutions more quickly. Simultaneously, operators of the platform gain better control over architecture, costs, and information assets. By centrally delivering common services, it becomes easier for Karolinska to steer and adapt solutions to new requirements. We don't need to customize each individual application, which is an advantage, as all applications require some form of integration with, for example, common security services.

The platform is designed to meet requirements of correctness, availability, and confidentiality. By offering central services using a "as a service"-model, we can also share costs and manage capacity

planning across applications. Since substantial parts of the hospital's operations constitute critical societal functions, we have chosen a hybrid model for the platform. The main services will reside in the region's data centers, but we also complement with outsourced services when appropriate and cost-effective.

An important aspect of this strategy is the need to develop our own development capability. The "as a service"-model forces us to build in-house expertise in operating and developing the technology platform, and to be able to develop applications ourselves. By building our own capability, we make sure that the platform is usable and that we can quickly implement new needs with our own resources.

Karolinska will benefit indirectly through improved information management, data governance and a more modern and standardized IT environment. Direct benefits are realized by developing and implementing applications on the platform. Development can be carried out by Karolinska itself and/or by external parties. Through what we call the Digital Health Platform, we have a strategy for how applications can be rapidly and securely implemented in a more agile and efficient way compared to traditional system development. Figure 1 show an overview of the platform services which are described in more detail below.

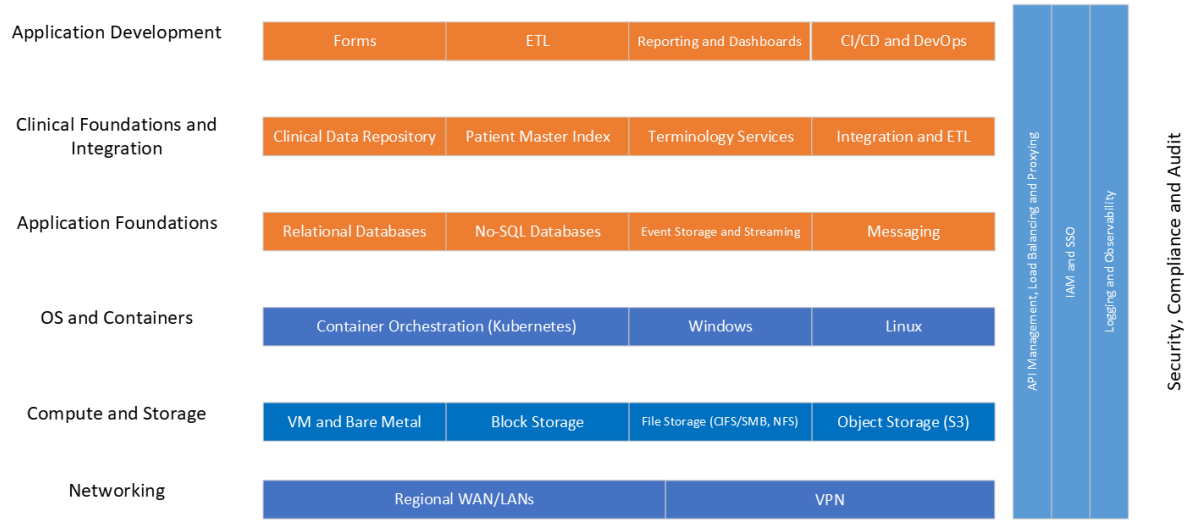

Figure 1 Visualization of services of Karolinska's Digital Health Platform

4 Current Status

Karolinska University Hospital is a part of Region Stockholm and largely utilizes shared IT services provided by the Region Stockholm's Service Management Department (SF). Since the end of 2022, a significant portion of SF's operations has been outsourced to an external vendor through a contracted agreement and Karolinska accesses these services through the Service Management Department. The agreement encompasses three types of services:

1. IT-operation within the region's existing data centers.
2. IT-operation within the vendor's private cloud environment.
3. IT-operation within public cloud environments.

These services will be provided following an implementation project[1]. In addition to SF's services, Karolinska, as the only hospital in the region, also has its own IT infrastructure. This infrastructure uses data center services from SF, such as electricity, cooling, perimeter security, and networking. It is supplemented with hardware, licenses, and IT-operations using Karolinska's own personnel.

Today, the Digital health platform is entirely based on Karolinska's own It-services, as SF's services under the new agreement have not been fully implemented. However, the technological choices are standardized, enabling a future migration when appropriate.

5 Platform Infrastructure in Karolinska's own IT-department

Compute, and storage services are based on technology from primarily VMware, IBM, and Red Hat running on HPE hardware. The service offerings include:

- Virtual machines (Windows, Linux) using VMware ESX/vCenter.
- Several Red Hat OpenShift Container Platform (OCP) clusters for container orchestration based on Kubernetes.
- Shared storage (block, file, object) through SAN (iSCSI, FibreChannel) and RedHat Ceph Storage.
- Backup using IBM Spectrum Scale Protect.

The OCP clusters we have today are based on both virtual machines and bare metal servers spread across two datacenters. Developers are provided with OCP projects and a set of common services that they can use to develop and deploy applications following a DevOps model. Common platform services, described below, are either based on OCP or virtual servers.

6 Application Foundations

The OCP (OpenShift Container Platform) serves as a solid foundation but needs to be complemented and/or augmented with other services to meet application and regulatory requirements. Services, such as logging, are replaced and services that aren't natively present in OCP, such as database hosting, are added. These services give us a comprehensive application and development platform, enabling the realization of a wide range of applications and solutions.

Continuous Integration and Continuous Deployment Processes (CI/CD)

We have developed custom CI/CD flows using the GitHub Enterprise Cloud SaaS service to help developers with source code management, build processes, and code review. Most builds, use local build agents[2] to allow access to our internal networks. Build artifacts and container images are stored in a central registry (JFrog Artifactory), where we also have the capability to perform security

[1] For internal reference : https://intranat.sll.se/stod-i-arbetet/organisation/projekt/inforandeprojekt---kapacitets-och-driftstjanster

[2] Github self-hosted runners

scans and flag vulnerabilities. We utilize built-in services within OCP (GitOps via ArgoCD[3]) for software automation and delivery to operational environments.

7 Security Services

The built-in logging services of OCP have been replaced with an external logging service (Splunk Enterprise) that enables us to manage logs for various purposes: technical system logs as well as compliance control (audit logs). Secrets are managed in an external vault service (HashiCorp Vault), allowing us to securely handle confidential information such as passwords and certificates in our GitOps workflows.

Authentication and authorization are implemented using regional services (SLL IdP and AD) using a federated Identity Provider (RedHat Single Sign-On) to comply with national and regional requirements[4]. Incoming traffic is handled either within OCP's built-in services or through a separate API security service (Kong Enterprise). The API security service is integrated with the Identity Provider and a policy engine (Open Policy Agent) for authorization control.

8 Integration and Data Storage

The Digital Health Platform offers a variety of integration and data storage services. These services are categorized as follows:

- Relational databases: PostgreSQL and Microsoft SQL Server
- Integration frameworks: Apache Camel and Camel K (Red Hat Application Foundations)
- Queue and event management: AMQP (Red Hat Application Foundations) and Apache Kafka (Confluent Platform)
- No-SQL databases: Couchbase

9 Clinical Foundations

The shared platform services provide a foundation for developing practically any type of application. To accelerate the development of applications for clinical needs, where we typically deal with healthcare information, we are building a Digital Health Platform. These foundational services provide storage and retrieval capabilities using three main components:

- A clinical data repository (CDR) based on international standards (openEHR and HL7 FHIR).
- Low-code development tooling such as form runtimes based on openEHR or BI reporting tools.
- A traditional enterprise data warehouse (EDW) for integrating non-clinical information.

By utilizing openEHR, we can reuse models developed regionally, nationally, and/or internationally. We also store information in an open, vendor-neutral format. Model specifications in openEHR format are machine-readable, enabling us to rapidly generate application components instead of wiring code. The development of applications based on the clinical data repository and the repository

[3] ArgoCD is an open-source solution integrated in OCP. Karolinska has the flexibility to adopt alternative solutions if needed.

[4] SITHS-card authentication and integration with the HSA-catalogue

itself leverages the shared services within the platform. The benefits and risks of openEHR and the clinical data repository are detailed in an external report commissioned by the region[5].

10 Cambio Health Platform – openEHR CDR

One of the challenges with a clinical data repository is that there are few implementations that also support Swedish legislation, specifically the Patient Data Act (PDL). Therefore, we have, since 2021, been collaborating with Cambio AB to develop a CDR solution that also includes components allowing us to implement logging, access controls, and consent management.

Since the openEHR information model separates demographic data from clinical documents, we can employ pseudonymization to maintain a required level of privacy and data protection. Cambio's solution includes services for managing patient identities, as well as integrations to regional authorization services. Cambio's solution also has a form engine, associated openEHR-specific informatics modeling tools, and a terminology server.

11 Associated Standards

The concept of the Digital Health Platform is based on openEHR but is also supplemented with services for other standards. The dominant standard for information exchange on the market today is HL7v2, and to some extent, the newer standard HL7 FHIR. Even though the information is managed and stored in the openEHR reference information model, we want to be able to exchange information using HL7-based protocols. This requires integration and mapping capabilities, as well as an internal ability to manage profiles. Alongside the implementation of openEHR and the associated CDR, we have been systematically working with HL7 FHIR, especially. FHIR has many useful components for realizing administrative workflows and will play a significant role in a future ecosystem. We currently use tools from Firely to manage FHIR profiles (simplifier.net and Forge).

12 Data Warehousing

Karolinska has extensive experience in working with traditional analysis for planning and business intelligence. Our traditional existing data repository, called "Karda," is currently used alongside a reporting engine to provide operational information to the entire hospital. We will still have a need for a traditional solution where we model information ourselves, as openEHR and FHIR do not cover all types of information objects (such as finance and HR data). Currently, Karda is based on Microsoft technology (SQL Server, SSIS) with operations hosted by SF. To simplify the development of reports and the ability to consume information, we currently use a reporting tool from Tableau (Tableau Server).

[5] RS 2022-0070 ”Utredning om effekterna av att välja OpenEHR”

13 Transitions to the new Digital Health Plattform

Karolinska is working with IT-transition to move from our current state to our vision of a more modern technical landscape.

Upgrades in our Technical Landscape

Karolinska will look into different technical improvements in today's IT-landscape, examples are:

- Consolidating and integrating the existing Data Warehouse solution with the new Digital Health Platform.
- Better solutions for handling unstructured data and more advanced statistical analysis.
- Exploring various options to achieve a cost-effective Enterprise Data Warehouse (EDW) solution.
- Implement a data catalog to efficiently manage data products and increase data governance.
- Implement advanced analysis techniques and utilize artificial intelligence (AI) to enhance insights and the value derived from our data.
- Enhance security measures and ensure compliance with regulations and requirements.

Manage Unstructured Data

In addition to the interoperability aspects covered with FHIR and openEHR, we also require comprehensive solutions for managing unstructured documents, such as PDF files. In this context, we intend to utilize IHE XDS-based standards.

14 Dependencies to other strategies and initiatives

Karolinska's Digital Health Platform initiative has dependencies to other projects in Region Stockholm.

- Replacement of Region Stockholm's Electronic Health Record (EHR) - TakeCare. This project will purchase and implement a new EHR and discontinue TakeCare.
- Region Stockholm has an on-going initiative for Master Data