DOI: 10.26314/GARR-Conf22-proceedings 16

# The hospital research network for the GATEKEEPER project: a case study

Francesco Ricciardi, Sergio Russo, Stella Grazia Pastore, Francesco Giuliani

**Abstract.** Artificial Intelligence and Big Data can exploit the value of the data coming from research. Data protection is a major issue when research data are related to people health status. One objective of the GATEKEEPER European project is to deliver a data federation framework and infrastructure to collect and give access to data about the health and well-being of European citizens. Within the project, our pilot was focused on the control of Type 2 Diabetes through the Artificial Intelligence techniques. This paper is focused on the description of our experience in the design and setting-up of a network infrastructure used to conduct the piloting activities, that works in parallel with the hospital enterprise network and satisfies data protection requirements

**Keywords.** Artificial Intelligence, Big Data, Well-being, Type-2 Diabetes, Hospital network

## Introduction

Nowadays data constitutes a great asset in research. Technologies like Artificial Intelligence and Big Data can exploit the value of the data coming from research, giving the possibility to produce new scientific insights by the analysis of a large amount of information. This is particularly true in the healthcare field where a large amount of information is available in digital format. Data protection is a key factor in all research disciplines, especially if scientific results can be transferred to the industry. The European landscape of healthcare data protection is strictly regulated by the General Data Protection Regulation (GDPR). Healthcare digital infrastructures attract the interest of cybercriminals. In 2020 there was an increment of 177% of the attacks against healthcare institutions (Critical Insight, 2021), while in 2021 72% of the attacks directed at healthcare infrastructures resulted in a data breach (Verizon, 2022). A high level of security measures is required to avoid cybersecurity accidents. This level of protection often conflicts with the flexibility needed in the manipulation of research data. Additionally, healthcare data in hospitals is employed for research purposes but perceived in many cases non-essential for the core business, leading to the choice of not exploiting the potential of the data not to increase the risks related to cybersecurity accidents.

The European Commission is actively funding research programs aimed at improving healthcare delivery across Europe. The GATEKEEPER project (Gatekeeper project, 2022) is an EU-funded initiative under the Horizon 2020 Framework Programme and its main objective is to enable the creation of a platform that connects healthcare providers, businesses, entrepreneurs, older adults and the communities they live in, to originate an open, trust-based arena for matching ideas, technologies, user needs and processes, aiming at ensuring healthier independent lives for the ageing populations. One of the objectives of




Palermo, 18-20 maggio 2022 - ISBN: 978-88-946629-1-7 - Creative Commons CC-BY

the project is to deliver a data federation framework and infrastructure to collect and give access to data about the health and well-being of European citizens. Within the GATEKEEPER project, Reference Use Case number 3 (RUC3) is focused on the control of Type 2 Diabetes through the AI techniques using both conventional data, extracted from clinical Electronic Medical Records (EMR) and Electronic Case Report Form (eCRF), and unconventional data gathered from wearable devices.
This paper is focused on the description of our experience in the design and setting-up of a network infrastructure to be used for research purposes that work in parallel with the hospital enterprise network. In particular, we will present a use case focused on the transfer of EMR data to the GATEKEEPER project data federation securely and in compliance with GDPR. We propose to use this setup as a reference model for hospital data sharing in research projects.

## 1. Objectives

Storing and moving data for research purposes is not a trivial task especially if personal data are involved. We designed a network infrastructure that could preserve information security.
The objectives of our design are:

- Compliance with GDPR rules, with the consent forms and research rules
- Guarantee of Confidentiality, Integrity and Availability of research and data from the HIS (Hospital Information System)
- Standard Interoperability
- Use of open-source, free solutions, when available
- Scalability
- Give the researcher the freedom of conducting research without sacrificing security requirements

## 2. Methods

The piloting activities of the GATEKEEPER project required the collection of clinical information and laboratory test results from blood exams. These data were collected in the EMR (Electronic Medical Record) and eCRF (electronic Case Report Form) three times during an enrolment period of 1 year (patient visits at baseline, after 6 months and after 12 months) according to the research protocol. The above-mentioned information needed to be transferred to the project data federation infrastructure. In parallel, data coming from a smartwatch worn by the patient flows through the manufacturers' servers to the same destination.
To satisfy the needs of researchers and security requirements, the whole hospital network has been split into two subnetworks interconnected through a firewall, a Hospital Enterprise Network and a Hospital Research Network, as shown in Figure 1. The patients' EMR data are stored in the clinical Enterprise Resource Planning (ERP) system inside the Hospital Enterprise Network. Laboratory test results flow from the Laboratory Information Management System (LIMS) to the clinical ERP.

The Hospital Research Network contains all the research related assets (physical servers, applications, databases, etc.). Within this network, we deployed the infrastructure for the conduction of the activities of the pilot. To collect the eCRF data, we developed a custom PHP web application integrated with the local Domain Controller for user authentication. This application is interfaced on one side with the Clinical ERP to collect the correct patient registry data to ensure data consistency all over the process. Registry data are then stored in a local database (PostgreSQL) within the Hospital Research Network. A registry data copy is necessary to allow the researchers to find the patient during the follow-up of the project. The application is in charge of automatically associating a project patient identifier that will be used at the project level for pseudonymization purposes. The same identifier is used on the smartphone app that collects and transfers the smartwatch data. Researchers collect eCRF data at each visit and create, at the same time, a record for EMR data extraction through this application.

A MIRTH connect (NextGen, 2022) middleware is in charge of extracting laboratory test results, once ready, from the clinical ERP. The values extracted are saved in the EMR data record created by the web application. Once all the data are ready, the MIRTH middleware transfers information to the project data federation through a REST web service. Data are associated with the project identifier. The RESTful interface accepts a customized JSON format and converts all the data according to the FHIR representation used at the project level. The FHIR data are then stored in the GATEKEEPER data federation platform. The data coming from the smartwatch are also converted in FHIR and stored on the same platform.

## 3. Results

The use of a custom architecture allows the conduction of the GATEKEEPER project activities with a minimum impact on the hospital enterprise network and ordinary hospital data flows. The use of separate resources and of a local database for research purposes accelerated the set-up of a server for the deployment of the project web application while reduces the attack surface of the hospital enterprise network.

The communication architecture doesn't require opening the access to the external networks, since the transmission of data is triggered internally by the MIRTH middleware.

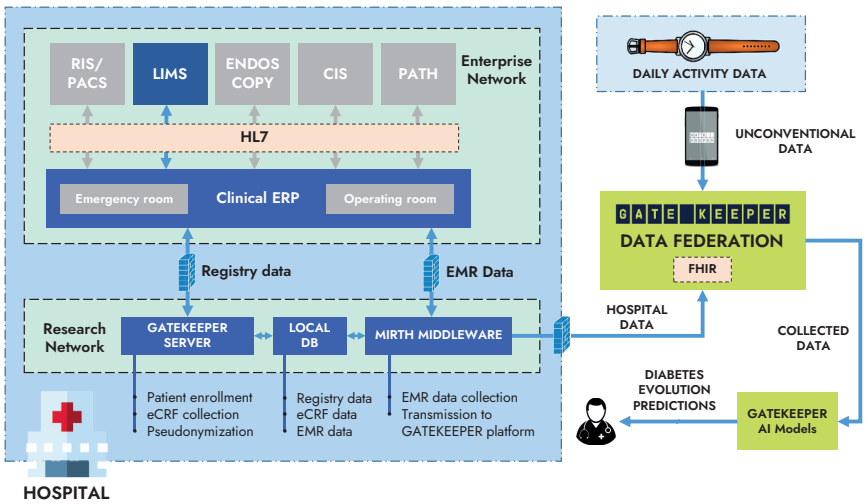


Fig. 1
GATEKEEPER project data flow

No services are exposed to the Internet, thus reducing the possibility of attacks. Access to the web application is granted through the least privilege principle while ensuring the user authentication by the network Domain Controller.

## 4. Conclusions

We presented a pilot experience focused on one possible use and configuration of the Hospital Research Network in the frame of the GATEKEEPER research project. The use of a separate network for research purposes has many advantages and a few disadvantages. First of all, the use of a separate infrastructure gives the researchers the flexibility to manage their resources, without undergoing the rigid requirements of the hospital cybersecurity rules. The set-up of a Hospital Research Network ensures the replication of the only information required for research purposes, in conformance with the GDPR data minimization principle. The replication of the data ensures also that researchers don't need to access the Hospital Enterprise Network, thus reducing the risks associated with data integrity and availability in the Enterprise Environment. The use of a separate network has the disadvantage of increased costs for infrastructure management, maintenance, etc. and requires a correct network configuration (at the firewall level) every time new data needs to be accessed in the enterprise network but this reduces the probability of errors and misconfigured rules.

Following our experience with the Gatekeeper project, we can conclude that, apart from the underlying infrastructure, considering data flows is of the utmost importance to ensure data protection. Equally important is to put in place appropriate pseudonymization techniques when sharing data outside the hospital borders.

## Acknowledgements

GATEKEEPER project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N° 857223.

## Bibliography

Critical Insight, «Healthcare Breach Report,» 2021.

Gatekeeper project, «GATEKEEPER PROJECT,» [Online]. Available: https://www.gatekeeper-project.eu/, 2022.

NextGen, «MIRTH CONNECT,» [Online]. Available: https://www.nextgen.com/products-and-services/integration-engine.

Verizon, «2021 Data Breach Investigations Report,» 2022.

## Authors

**Francesco Ricciardi** f.ricciardi@operapadrepio.it

Francesco Ricciardi (M) has a degree in Electronic Engineering and a PhD in Information Engineering. His research interests are in the field of Virtual and Augmented Reality applied to the healthcare field, assistive robotics and active, healthy ageing and cybersecurity. He is currently involved in fund raising, proposal writing, pilot execution and project management of European

research projects. He is author and co-author of some scientific papers appeared in many international journals.

**Sergio Russo** s.russo@operapadrepio.it
Sergio Russo (M) is a research engineer in the Innovation and Research Unit at the Research Hospital Casa Sollievo della Sofferenza. He has a MSc degree in Computer Science Engineering. He is currently involved in the digitalization and reengineering of processes and in project management activities for regional, national and european research projects, with a major interest in Assistive Robotics, AAL, IoT, orchestration and distributed systems. He is co-author of some scientific papers.

**Stella Grazia Pastore** sg.pastore@operapadrepio.it
Stella Pastore (F) has a degree in Chemistry and a PhD in Chemical and Molecular Science. She currently works at Fondazione Casa Sollievo della Sofferenza-IRCCS as Clinical Reseacher, supporting digital clinical trials, European funded pilot studies and project management activities. With a background in both academia and industry, her main interests are focused on healthcare and IoT world.

**Francesco Giuliani** f.giuliani@operapadrepio.it
Francesco Giuliani (M) is the head of the Innovation and Research unit of Casa Sollievo della Sofferenza Hospital. Thanks to the participation to national and international research projects, the unit conceives and conducts research and innovation projects for the clinical world. The area of interests are mHealth and eHealth, assistive robotics, Artificial Intelligence, Ambient Assisted Living, Scientometrics.