Supported by:

Federal Ministry of Health

on the basis of a decision by the German Bundestag

ELNET

GIHF AI

German Israeli Health Forum for Artificial Intelligence

Report "Technology and Security"

Data Access and Data Linkage: More data, better data, more connected data - but how?

This report is a summary of the first German Israeli Health Forum for Artificial Intelligence (GIHF-AI) Digital Health Roundtable-Workshop on April 7, 2022 and addresses health data access and linkage with a focus on data interoperability, cybersecurity, privacy, and regulation. It complements the GIHF-AI Policy Briefing "Innovation through interoperable linked health data and collaboration in science" from April 1, 2022. The video recording of the roundtable can be found on the ELNET YouTube Channel.

RECOMMENDATIONS FOR ACTION

- **The introduction of interoperable data standards such as LOINC, SNOMED and FHIR in the public and private sectors should be a top priority to guarantee the usability of data.** This is because the meaningful use of AI in healthcare depends on the quality, standardization, and quantity of available medical data. **Political regulation and the creation of incentives** can accelerate the implementation of data standards in the public and non-public sectors; the experience of the Israeli government has shown this. In addition, this can and should avoid the creation of parallel structures.

- **Health data protection should follow a risk-benefit approach, be sector-specific and be organized in a centralized manner.** The large number of cross-sectoral data protection authorities in Germany leads to dilution and unmanageable as well as unnecessary regulation, which not only creates intransparency, but in many cases is a concrete hindrance to research and science. Project-based and risk-based evaluation, based on the GDPR, minimizes data protection errors while enabling (collaborative) research. In addition, a **cybersecurity framework in the sense of Protect, Detect, Restore** should be applied to minimize data protection risks from cyberattacks.

- **The Federal Government of Germany should create and drive the appropriate infrastructure to accelerate the development process of a European Health Data Space. This also includes the rapid and widespread introduction of the new ePA 2.0,** including implementation and support for users (doctors and patients) and communication. This is because the creation of a European Health Data Space and further support for national initiatives such as the Health Data Lab (HDL) at the BfArM will ensure efficient exchange of and direct access to various health data in health care, health research and health policy.

RECOMMENDATIONS FOR ACTION

- To enable cross-institutional scientific collaboration, including in an international context, the **use of Privacy Preserving Record Linkage (PPRL) and Privacy Enhancing Technologies (PET)** should be considered. Where necessary, appropriate **regulatory adjustments** should take place. These methods offer researchers, with the help of specialized service providers, the ability to link data sets securely and in a privacy compliant manner.

- **Health data should be understood as a public good.** The use of FAIR (= findability, accessibility, interoperability, reusability) health data, while respecting the GDPR, can create the needed trust of the population in the use of their data. A **corresponding communication campaign** to educate and inform is recommended in addition to regulatory alignment.

- **The promotion and funding of a bilateral research project in the form of a call for proposals for a digital AI-based health application is recommended.** This is because Germany currently does not have the health data it needs and can therefore benefit from Israel's wealth of data, especially in the field of AI. Such a bilateral research project can act as a use pilot program for further research projects. The **call for proposals should address both public and non-public institutions** to achieve the best possible outcome. Israel's Ministry of Health has already been able to launch numerous successful projects with such tenders.[1]

International collaboration in digital health: interoperability is key

International scientific cooperation in the field of medicine depends on interoperable data, as recently demonstrated by the Corona pandemic. Exemplary for this is the standard data set for COVID-19 patients, the "German Corona Consensus Dataset" (GECCO). The dataset provides the scientific community with a common language and working basis.

It was used, among other things, in the Corona-Warn-App and the international research project ORCHESTRA, where data from all over the world were combined and harmonized to answer questions about indications, side effects and drug interactions. In addition, high-quality, standardized data is needed to validate AI analyses. Higher validity, in turn, leads to greater confidence in AI-based systems.

Furthermore, a healthcare system digitized by interoperable data facilitates medical documentation, mitigates communication-related treatment errors, and eases the flow of information. Institutions such as the German Institute for Standards (DIN) and, in the international context, the Joint Initiative Council for Global Health Informatics Standardization (JIC), with around 300 standards, are responsible for standardizing health data. These include LOINC, SNOMED and HL7, as well as the Global Digital Health Partnership (GDHP) under the umbrella of the WHO, or the Global Alliance for Genomics and Health (GA4GH) in the field of genomic data.

The institutions are committed to constant flow of health data, which requires, among other things, the use of an International Patient Summary (IPS). It contains essential health information as specified in EN 17269 and ISO/DIS 27269 and is designed to support, but not limited to, the unplanned, cross-border care application scenario.

In Germany, the electronic patient record (ePA) was introduced in 2022. Expected in 2023, ePA 2.0. will be even more patient-centric (patients will decide independently how their data may be used), have enhanced functionality, and most importantly, be

interoperable as it is based on FHIR (Fast Healthcare Interoperability Resources).

The healthcare data exchange standard developed by Health Level 7 (HL7) is considered the current "gold standard" and is increasingly used in an international context. The Israeli Ministry of Health is working at full speed to introduce FHIR standards across the board in the public and non-public sectors. The implementation of FHIR in Germany will be accompanied by facilitation for collaboration.

Nevertheless, there are challenges, such as the integration of results from DIGA (Digital Health Application) or MIO (Medical Information Objects) as well as the feed of radiological findings and AI-based data. The Interop Council for digital health in Germany, consisting of more than 100 experts, has been coordinating processes around data interoperability since December 2021 and will address these.

In the area of research, the Medical Informatics Initiative (MII) of the 34 German university hospitals, which is funded by the German Federal Ministry of Education and Research (BMBF), is primarily concerned with collaboration on medical data. As with all cross-institutional collaborations, the challenges include, in particular, maintaining the privacy of sensitive patient data. For this reason, federated learning techniques are used. Here, AI calculations that are important for training the algorithm are performed directly on the end device. With these methods, the data does not leave the facilities.

In addition, standardized data sets such as GECCO should be used. The LOINC (Logical Observation Identifiers Names and Codes) terminology, which has existed for 15 years and contains 90,000 terms, is part of the data set. Furthermore, SNOMED CT is used, the most comprehensive terminology with 340,000 medical terms. Germany joined the SNOMED community in 2021. By comparison, Israel joined the SNOMED community as early as 2012 and therefore has valuable experience in data standardization.

Additively, the ISO standard "Identification of Medical Products" (IDMP) is increasingly being used at university hospitals such as the Charité – Universitätsmedizin Berlin. Further data standards in the area of pharmaceuticals are to be made available (soon) by the Federal Institute for Drugs and Medical Devices (BfArM). Private software companies, on the other hand, often use their own standards, which are a hindrance in the context of interoperability. Expanded regulation of data standards would remedy this problem.

For the international statistical classification of diseases and health problems, both Israel and Germany use the International Classification of Diseases (ICD) standard specified by the World Health Organization (WHO). The current ICD-10 is used in 100 countries and has over 11,400 codes. In both Germany and Israel, work is underway to introduce the new, improved ICD-11 across the country, which came into force in 2022.[2]

Protecting patient data: privacy and cybersecurity in healthcare

Data protection and cybersecurity should be a top priority, especially in the healthcare sector, as it involves highly sensitive data. In order to build trust in digitized applications, guarantee the smooth running of clinics, practices and research institutions, and serve research, healthcare data must follow the highest data protection and cybersecurity standards. Nevertheless, it is important to keep in mind that innovative and timely science and patient care depend on data being allowed to be used and shared. A risk-based approach including quantitative and qualitative risk evaluation prior to each project can prevent overregulation from stifling scientific progress.

At the same time, risks are reduced to a minimum, because sensible risk-benefit prioritization is the key to privacy-compliant science. It should also be noted that the European Union (EU) General Data Protection Regulation, on which Israeli data protection regulation is also based, provides multiple legal bases in the area of data protection. Different guidelines apply depending on whether the data

are to be used for outpatient or inpatient medical care, research, or the creation of patient files. It is often assumed that explicit patient consent is the only legally possible basis, although there are a total of six legal bases in the GDPR for the use of health data. These opportunities for adaptation should be exploited.[3]

It is also important to apply the highest standards in the area of cybersecurity. After all, cyberattacks on medical institutions can bring entire hospitals to a standstill as well as cause major data leaks. Typically, these attacks take place in the institution's Active Directory, the database. A risk-based backup of the Active Directory in the sense of Protect, Detect and Restore therefore leads to additional security and should be implemented across the board. Specialized service providers can help to equip the systems accordingly.[4]

Storing and using health data: Israel and Germany

In Israel, health data has been stored digitally since the 1990s and has therefore been available in completely digital form for over twenty years. The healthcare system is based on the four health insurance companies (Kupot Holim), of which Clalit is the largest. In addition to hospitals, outpatient clinics and pharmacies, they also maintain innovation centers. These centers are responsible for the data of all people insured with the respective health insurance company, who as a rule remain with one health insurance company from birth to death. The data can be used both internally and by external parties in the form of cloud solutions. Interoperable data standards ensure maximum usability, also in the context of international scientific cooperation. By anonymizing the patient data, the data can be made available to researchers without risking data protection errors. In the course of the corona pandemic, the whole world benefited from this wealth of data, which massively facilitated vaccine development.[5]

The German health care system is currently working hard to build a health data ecosystem so that it can benefit from its own enormous wealth of data. The largest nationwide initiative is the Health Data Lab (HDL), Forschungsdatenzentrum Gesundheit (FDZ) in German, at the BfArM, which is currently being set up to provide researchers in particular with patient data for scientific purposes. Data from around 72 million people with statutory health insurance since 2009 are to be made available for research in the future, in particular to health insurance companies, universities and scientific institutions, but not to the private sector. Data from ePA 2.0., DIGA and MIO will provide content to the HDL.[6]

The evolution of healthcare through data linkage

Scientific cooperation in the national and international field depends on data cooperation. The use of large data sets is essential for both the validation of results and the generation of new hypotheses. If an institution does not have enough data, linkage of data from other institutions is required. Using Privacy Preserving Record Linkage (PPRL) and Privacy Enhancing Technologies (PET), which are technologies already successfully in use in Israel's healthcare system today, institutions conduct research together without ever seeing the collaborating partner's data. Data collaboration takes place, for example, through Data Encryption, Synthesis, Differential Privacy (DP), Trusted Execution Environments (TEE), and Federated Learning (FL). In addition, there are software and technologies that have built-in verification mechanisms to avoid sharing sensitive data. A warning message alerts users that there is a privacy issue. In order to ensure interoperability, the data is harmonized to, for example, FHIR standard.[7]

Bibliography

1. **ELNET Deutschland:** "GIHF-AI Digital Health Roundtable Data Access & Data Linkage", 07.04.2022 in *https://www.youtube.com/watch?v=7MrHyhXmcjQ;* **ELNET Deutschland:** "GIHF-AI Digital Health Workshop Data Access & Data Linkage, 07.04.2022; **ELNET Deutschland:** "Innovation durch interoperable, verknüpfte Gesundheitsdaten und Wissenschaftskooperation: GIHF-AI Policy Briefing „Technik & Sicherheit", 01.04.2022 in *https://gihf-ai.eu/wp-content/uploads/2022/04/GIHF-AI_PolicyPaper_TechSec_Data-4.pdf;* **ELNET Deutschland:** "Digitalisierung und Innovation im deutschen und israelischen Gesundheitswesen: Eine Bestandsaufnahme", 03.03.2022 in *https://gihf-ai.eu/wp-content/uploads/2022/03/220301_GIHFAIBestandsaufnahme-1.pdf.*

2. **Thun, Prof. Dr. Sylvia:** "GIHF-AI Digital Health Roundtable Data Access & Data Linkage", 07.04.2022 in *https://www.youtube.com/watch?v=7MrHyhXmcjQ;* **ELNET Deutschland:** "Innovation durch interoperable, verknüpfte Gesundheitsdaten und Wissenschaftskooperation: GIHF-AI Policy Briefing „Technik & Sicherheit", 01.04.2022 in *https://gihf-ai.eu/wp-content/uploads/2022/04/GIHF-AI_PolicyPaper_TechSec_Data-4.pdf;* **theblue.ai:** "Federated Learning / Föderales Lernen", 02.09.2019 in *https://theblue.ai/blog-de/federated-learning-foederales-lernen/.*

3. **Shmerling Magazanik, Limor** "GIHF-AI Digital Health Roundtable Data Access & Data Linkage", 07.04.2022 in *https://www.youtube.com/watch?v=7MrHyhXmcjQ.*

4. **Keizers, Oliver:** "GIHF-AI Digital Health Roundtable Data Access & Data Linkage", 07.04.2022 in *https://www.youtube.com/watch?v=7MrHyhXmcjQ.*

5. **Pomerance, Yael:** "GIHF-AI Digital Health Roundtable Data Access & Data Linkage", 07.04.2022 in *https://www.youtube.com/watch?v=7MrHyhXmcjQ.*

6. **Schneider, Dr. Katharina:** "GIHF-AI Digital Health Roundtable Data Access & Data Linkage", 07.04.2022 in *https://www.youtube.com/watch?v=7MrHyhXmcjQ.*

7. **Lalazar, Dr. Gadi:** "GIHF-AI Digital Health Roundtable Data Access & Data Linkage", 07.04.2022 in *https://www.youtube.com/watch?v=7MrHyhXmcjQ;* **Shainski, Rina:** "GIHF-AI Digital Health Roundtable Data Access & Data Linkage", 07.04.2022 in *https://www.youtube.com/watch?v=7MrHyhXmcjQ.*

Supported by:

Federal Ministry of Health

on the basis of a decision by the German Bundestag

ELNET

GIHF AI

German Israeli Health Forum for Artificial Intelligence

ELNET Deutschland e.V.

Albrechtstraße 22
10117 Berlin
deutschland@elnetwork.eu

elnet-deutschland.de

@ElnetD

@ELNETDeutschland

GIHF-AI
German Israeli Health Forum for Artificial Intelligence

Albrechtstraße 22
10117 Berlin
info@gihf-ai.eu

gihf-ai.eu

@GIHFAI

CONTACT

Carsten Ovens

Executive Director
ELNET Deutschland

Lea Ledwon

Program Manager GIHF-AI
ELNET Deutschland

Author: Lea Ledwon
Date: 10.05.2022