ICT Solutions for Brilliant Minds

CSC

Position Paper 8.7.2021
European Health Data Space
CSC – IT Center for Science Ltd.

Due to CSC's vast experience with offering sensitive data services for research[1], we have been following the plans for a European Health Data Space with curiosity and believe we have a lot of expertise and ideas to contribute to the process. Considering our specialisation in services for research, we are commenting the plans in this paper and in our answers to the questionnaire from the point of view of secondary use of health data, with the benefit of European citizens as the ultimate target. Well-managed, accessible and interoperable research data and the ability to process it, is a key prerequisite for the development of healthcare on the primary side.

--

CSC agrees with the need to pay special attention to the issues and barriers that are specific to the exchange of and access to health data and therefore welcomes the Commission's intention to create a dedicated legislative framework for the European Health Data Space. However, it must be ensured that the sector-specific data spaces and their respective legislative frameworks do not become silos but are developed in coherence, making them interoperable and thus supporting the ultimate goal of creating one unified European data space. This is necessary for enabling innovation in the health sector.

Building infrastructures and tools for managing and processing health data is crucial for the wellbeing of European citizens, and it is of utmost importance to ensure the ability for researchers to access and use data across borders and institutions, to enable research and innovation that aims at more efficient treatments, personalised medicine and better cure for rare diseases. The EU is currently making remarkable investments in data management and high-performance computing capacities (e.g. EuroHPC[2]), which will boost the research and innovation capacity also in the health field. It is important to fully leverage these investments and capacities.

The EHDS is likely to be a federated infrastructure that connects existing data repositories rather than a new, centralised one. This is a good approach as it builds on previous investments and improves information security by allowing the data to stay in its original location even while accessed and processed by users from other countries. It must be noted, however, that a federated model requires significant efforts to ensure the interoperability between data in the separate repositories. This implies creating common data management practices as well as making substantial investments in the practical implementation of the interoperability requirements.

[1] https://research.csc.fi/sensitive-data
[2] https://eurohpc-ju.europa.eu/

When creating the common data management practices, coherence across the European data spaces and all data related regulation (such as the upcoming Data Governance Act[3]) must be ensured. This requires basing the regulatory framework of the health data space on the FAIR principles[4] and the European Interoperability Framework[5]. All four layers of interoperability must be addressed, including legal interoperability that will require making sure that any existing national legislation related to the scope of the EHDS legislation will be brought in line with the latter. Lessons learnt from similar legislative processes at the national level must also be taken into account. For example, the experience from Finland's Act on the Secondary Use of Health and Social Data[6] highlights the need to ensure a sufficient transition period and resources to those implementing the new policy.

When creating a data space dedicated to health data, particular attention must be paid to developing the processing of sensitive data in accordance with data protection rules and the MyData principles[7] aiming to empower individuals by improving their right to self-determination regarding their personal data ('data sovereignty'[8]). A right balance between guaranteeing data protection and encouraging data altruism must be struck, and policies related to these two concepts developed in a coordinated way.

It must be ensured, that sensitive data can be used for the purposes of research and innovation, bearing in mind that there are technical means for securing anonymity and privacy of individuals. It is recommendable to use federated authentication and authorisation frameworks, such as the ELIXIR AAI[9], based on global GA4GH standards[10]. Such technologies secure full control to all data access by using authorised Data Access Committees. All non-authorised access to sensitive data can be thus disabled.

Another key issue when dealing with health data is to ensure that European health data is stored and processed in secure locations and never outside the European Economic Area, unless the third country in question is legally committed to similar data management rules and principles as those applied in Europe. It must be kept in mind that the creation of a European data space does not require storing data in one place or even moving it anywhere. In a federated model, data can be accessed and processed in its original location, having connected existing infrastructures, made them interoperable and filled the gaps they may have. A first step of data federation is to ensure that metadata is interoperable.

The European Health Data Space must be run by operators possessing excellent ICT competences and built making use of existing projects, practices, policies and infrastructures related to cross-border use of health data. Examples of ongoing efforts to be leveraged include

[3] https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52020PC0767
[4] https://www.go-fair.org/fair-principles/
[5] https://ec.europa.eu/isa2/eif_en
[6] https://stm.fi/en/secondary-use-of-health-and-social-data
[7] https://mydata.org/declaration/
[8] https://datasovereigntynow.org/our-mission/
[9] https://elixir-europe.org/services/compute/aai
[10] https://www.ga4gh.org/genomic-data-toolkit/data-security-toolkit/

the TEHDAS[11], Beyond 1 Million Genomes (B1MG)[12], INTERVENE[13] and CINECA[14] projects as well as the work done in the framework of the Research Data Alliance[15], European Open Science Cloud[16] and GAIA-X[17]. In particular, the principles created in the B1MG project for Ethical, Legal and Social Issues (ELSI) as well as for federation are useful starting points for further development.

Digital infrastructures form an entity that must be developed in convergence, aiming at synergetic data ecosystems, which can pave the way for world-class research and innovation. This means for example, that linkages of the health data space with high-performance computing capacities and high-speed digital connections must be ensured. CSC is pleased to note that the Commission plans to also address digital health products and services, including AI systems, in the health data space legislation. This must be prepared in coherence with the upcoming European legal framework for AI[18].

Kimmo Koski
Managing Director
kimmo.koski@csc.fi

Irina Kupiainen
Public Affairs Director
irina.kupiainen@csc.fi

[11] https://tehdas.eu/project/
[12] https://b1mg-project.eu/
[13] https://www.interveneproject.eu/
[14] https://www.cineca-project.eu/
[15] https://rd-alliance.org/
[16] https://ec.europa.eu/info/research-and-innovation/strategy/strategy-2020-2024/our-digital-future/open-science/european-open-science-cloud-eosc_en
[17] https://www.gaia-x.eu/
[18] https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A52021PC0206