

*Disclaimer: This content is an extract from a draft version of the Horizon Europe Health Cluster work programme*

## **HORIZON-HLTH-2022-IND-13-02: Scaling up multi-party computation, data anonymisation techniques, and synthetic data generation**

**Expected Outcomes:** *This topic aims at supporting activities that are enabling or contributing to one or several expected impacts of destination 6 “Maintaining an innovative, sustainable and globally competitive health industry”. To that end, proposals under this topic should aim for delivering results that are directed, tailored towards and contributing to all of the following expected outcomes:*

- The EU contributes strongly to global standards for health data through enhancement of common European standards for health data (including medical imaging data) by researchers and innovators. Researchers and innovators contribute to GDPR compliant guidelines and rules for data anonymization.
- Innovators have access to advanced secure data processing tools to test and develop robust data-driven digital solutions and services in response to the needs of researchers, clinicians and health systems at large.
- Cross-border health data hubs further facilitate the innovation process by providing secure, trustable testing environments for innovators.
- Clinicians, patients and individuals use a larger variety of high quality data tools and services for wellbeing, prevention, diagnosis, treatment and follow-up of care.
- Researchers and innovators have more opportunities for testing and developing GDPR compliant data driven solutions based on actual needs of the health care environments.

**Scope:** It is essential to speed up and facilitate innovations in the field of data-driven tools and services for wellbeing, prevention, diagnosis, treatment and follow-up of care, among others. However, limited access by developers to health data and secure testing environments hinder the development of innovative data-driven digital health products and services.

Therefore, the proposals are expected to scale up multi-party computation, data anonymisation techniques and synthetic data generation. To ensure privacy, the data analytics should be conducted in a distributed way among processors that grant third parties access to analysis outcomes but not to the underlying data. The developers should have access to distributed testing data sources at large scale, with a view to improving the speed and robustness of multi-party computation solutions for innovators. The aim is to allow secure GDPR-compliant data processing for research, and clinical purposes.

The proposals should consider the use of synthetic, i.e. artificially generated, data as they allow researchers and developers to test, verify and fine-tune algorithms in large-scale data experimentations without re-identifiable personal data.

In addition, the proposed anonymisation techniques will have to be sophisticated and robust enough to tackle the challenge of anonymised data sets that still make it possible to trace back to individuals.

The proposals are expected to foster the development of secure, interoperable, transparent - and therefore trustable - cross-border health data hubs that can facilitate the provision of the required testing environments for innovators. This will support the uptake of new data tools, technologies and digital solutions for health care.

To this end, integration of national/regional health data hubs/repositories/research infrastructures is appropriate to achieve the scope of the topic.

***The proposals are expected to address all of the following areas:***

- Consolidate and scale up multi-party computation and data anonymisation techniques and synthetic data generation to support health technology providers, in particular SMEs.
- Support the development of innovative unbiased AI based and distributed tools, technologies and digital solutions for the benefit of researchers, patients and providers of health services, while maintaining a high level of data privacy.
- Advance the state-of-the-art of de-identification techniques, to tackle the challenge of anonymised datasets that can be traced back to individuals.
- Develop innovative anonymisation techniques demonstrating that effective data quality and usefulness can be preserved without compromising privacy.
- Explore and develop further the techniques of creating synthetic data, also dynamically on demand for specific use cases.
- Widen the basis for GDPR-compliant research and innovation on health data.
- Ensure wide uptake and scalability of the methodologies and tools developed, promote high standards of transparency and openness, going well beyond documentation and extending to aspects such as assumptions, architecture, code and any underlying data.

In this topic the integration of the gender dimension (sex and gender analysis) in research and innovation content is not a mandatory requirement.