

# **HARMONY Anonymization Concept Reconciles Data Quality, Safety, and Privacy**

**Ensures careful, judicious, and uncompromising adherence to ethical guidelines and the applicable Data Protection (DP) legislation**

**Data has the enormous potential to improve the health of patients around the world. Artificial intelligence (AI) systems analyze huge amounts of data, transforming it into valuable clinical knowledge and predictive models. To ensure that people trust technology, responsible privacy and security guidelines must be developed. How has the HARMONY Alliance implemented ethical and legislative principles for AI and Big Data?**

**HARMONY's Big Data research activities carry significant potential for improving cancer patients' treatment options. Still, we take great care to follow ethical principles for biomedical research to ensure that no mistakes are made. One example is the question of personal data security. Moreover, we must ascertain the quality of the data and their analysis, as only sound science is ethically justifiable.**

## **More than Data Safety and Privacy**

Everyone involved with the HARMONY Alliance understood that data security and privacy would always be the core of our data platform architecture. The aim of HARMONY has been ambitious from the very beginning: to establish a pan-European repository of longitudinal health data related to 7 hematological malignancies, as well as an additional database for studying childhood leukemias and lymphomas. If we want to help patients, we have to make sure they trust our technology. That means that uniting data from various diverse sources and ensuring that they are interoperable and reliable is not only a technological challenge, but a societal one as well. That means that we have to handle databases with special care. How do we - in the HARMONY Work Packages and HARMONY Research Projects - process data for the benefit of patients and in a way that guarantees that an ethical methodology is used for Big Data analysis?

At the HARMONY Big Data Platform, we not only have scientific ambitions, but we seek to set some of the highest standards in AI and data ethics. Our first step was to establish core principles for data safety, reliability, security, privacy, and anonymity, which included regulations that were already in force. Next, we moved past considering legislative frameworks and created an internal ethical code that articulated a clear vision of what we wanted to achieve, setting out rules that bound all the stakeholders and were relevant at every stage of the project. All the data processing guidelines also included the principles of fairness, transparency, and accountability.



*“We work closely with HARMONY’s medical, IT and big data experts to see that our data processing solution both provides maximum protection for data donors and leaves their data rich enough for valid scientific analysis. The elaborate data processing concept we developed was put under extensive scrutiny by external legal and ethics experts”,* says **Christiane Druml, UNESCO Chair in Bioethics at MediUni Wien, Director of the Josephinum, Medizinische Universitaet – Wien** and former HARMONY Work Package 8 Leader.

## Responsible Data Processing

The database of the HARMONY Big Data Platform include complex information on Hematologic Malignancies (HMs), including diagnostics, genomics, treatment choices, and outcomes, thus allowing for the processing and analysis of diverse data originally collected by hospitals, national study groups, clinical trials, and other groups. In this context, the HARMONY Alliance ensures careful, judicious, and uncompromising adherence to ethical guidelines and the applicable Data Protection (DP) legislation, i.e. the EU’s General Data Protection Regulation (GDPR) and local DP legislation.

For prospective data collected in studies performed under HARMONY’s oversight, the consortium requires the use of standard language in the patient’s informed consent form, which provides that explicit consent be given for the further use of data. For retrospective studies, however, an explicit informed consent form covering secondary data use within the consortium is not available and retrospective consent is not a feasible solution. Consequently, the Alliance has developed a modus operandi that will ensure that only adequately anonymized records are used and processed on the HARMONY Data Platform.

## The HARMONY Anonymization Concept

It has long been argued that anonymization (i.e. redacting data such that the data subject is not or is no longer identifiable) cannot be guaranteed without rendering data useless for medical research. However, anonymization in a legal sense does not require that data be redacted in a way that makes it entirely impossible (for example through legal or technical means) to identify the individual concerned. Rather, de-facto anonymization is sufficient in order to exclude the qualification of the relevant data as “personal data”; i.e. sufficient anonymity is ensured, as identification of the data subject would require an unreasonable amount of effort. In light of this, anonymization can be achieved through a combination of technical methods, such as suppression, generalization, and perturbation to the extent that any such effort does not compromise the scientific goals of the study; these techniques are supplemented by data access restriction and organizational security measures.

\* Regulation (EU) 2016/6791, the European Union’s new General Data Protection Regulation (GDPR), regulates the processing by an individual, a company or an organization of personal data relating to individuals in the EU. The regulation is an essential step to strengthen individuals’ fundamental rights in the digital age and facilitate business by clarifying rules for companies and public bodies in the digital single market.



## HARMONY Anonymization Concept Reconciles Data Quality, Safety, and Privacy

It is in this manner that we developed the HARMONY Anonymization Concept, which ensures that the intended import of data into the HARMONY Big Data Platform, as well as the subsequent use of such data as envisaged by the HARMONY Project, comply with ethical guidelines and all applicable data protection laws at the EU level, which include meeting the requirements of the General Data Protection Regulation (GDPR), without impacting the clinical value of the relevant data. The HARMONY Anonymization Concept takes into account all necessary factors to ensure that the case-by-case assessment of every single database is complete, and that no provisions required by applicable data protection laws are omitted.

Furthermore, HARMONY is also committed to regularly revisiting its anonymization protocols in light of new technical developments in order to ensure that the direct or indirect identification of individuals remains “impossible without unreasonable effort.”

*“The biggest issues with anonymization were two-fold: on the one hand, we couldn’t distort the data, which could have made sound scientific analysis impossible. On the other hand, the quality of technical anonymization procedures depends on the current state of technology – what is considered secure today might be easy to hack tomorrow. That’s why we went for our De-Facto-Anonymization concept solution, which comprises a two-step pseudonymization architecture complemented by extensive organizational measures, such as rigid data access control”,* **Christiane Druml** adds.

The HARMONY Ethics Advisory Board consists of Prof. Emeritus Peter Bauer, Prof. dr. Inez de Beaufort, Erasmus MC, The Netherlands and Prof. Federico de Montalvo Jääskeläinen, Universidad Pontificia Comillas | UPCO, Spain.

[www.harmony-alliance.eu](http://www.harmony-alliance.eu)

[Follow our LinkedIn Page](#) | [Follow us on Twitter](#) | [Subscribe to our Youtube channel](#)

**About the HARMONY Alliance:** a public-private European Network of Excellence for Big Data in Hematology, established in January 2017. Our mission is to unlock and spread valuable knowledge on hematologic malignancies (blood cancers) among a large number of stakeholders, with the goal to harness and mine Big Data to speed up the development of improved treatments for patients and more effective treatment strategies. HARMONY currently has 53 Partners and 32 Associated Members from 22 countries. HARMONY funded through the Innovative Medicines Initiative (IMI) (LINK), Europe's largest public-private initiative aiming to speed up the development of better and safer medicines for patients. HARMONY has received funding from IMI 2 Joint Undertaking and is listed under grant agreement No. 116026. This Joint Undertaking receives support from the European Union’s Horizon 2020 Research and Innovation Programme and the European Federation of Pharmaceutical Industries and Associations (EFPIA) (LINK). IMI supports collaborative research projects and builds networks of industrial and academic experts in order to boost pharmaceutical innovation in Europe.

