Strategic Focus Area
Personalized Health and Related Technologies

Final Scientific Report

General Information

Project number:	2017-201
Project title:	DPPH: Data Protection in Personalized Health
Reporting period:	01/04/2018 - 31/12/2021
PI:	Jean-Pierre Hubaux
Co-PIs:	Bryan Ford, Jacques Fellay, Effy Vayena, Olivier Verscheure

Part A - Individual report

Aims and milestones of the project (short)

The DPPH project aimed at addressing the main privacy, security, scalability, and ethical challenges of data sharing for enabling effective P4 medicine, by defining an optimal balance between usability, scalability and data protection, and by deploying an appropriate set of computing tools. The project was organized in three main phases. The first phase, carried out over its first year, was devoted to the collection of requirements and the development of an initial working prototype. During the second phase which lasted for two years, the core R&D activities were performed in order to advance the initial prototype and address the identified requirements, aligning the developments with the needs of SPHN. The third and final phase of the project focused on deployment, testing and validation of the developed prototype.

In more detail, after gathering the requirements in WP1, the project focused on producing an early prototype (WP2) that was moved into deployment on three Swiss University Hospitals through the companion MedCo project. WP3-WP7 focused on developing the tools comprising the DPPH security and privacy framework, which increased the functionality of the aforementioned prototype, while WP8 integrated these to the final DPPH platform. Finally, WP9 analyzed the ethical and legal framework and its impact on the developed tools.

Final status

The final status of DPPH (December 2021) has followed the original work plan, with some variations to account for the zero-cost project extension that was formally filed to and accepted by the PHRT administration.

The techniques we have developed in the framework of the project have reached a level of maturity and have generated so much interest (in the health data domain and beyond) that we have decided to launch a start-up, called Tune Insight (https://tuneinsight.com). We incorporated it in September 2021 and raised pre-seed capital from a Zurich-based VC, Wingman Ventures. As of January 2022, the company is fully operational and is taking care of the further development, the deployment, and the maintenance of some of the software tools developed in the framework of DPPH.

More details about the project's final status are provided in the following paragraphs, broken down according to the work packages that were active throughout its timeline:

WP1 (Requirements): The goal of WP1 was to gather information about security, privacy, ethical and functional requirements necessary to develop the DPPH software solutions that enable P4 medicine in Switzerland. Following up on the work done in the first two years of the project, in 2020 we wrapped up the main contributions for this work package and produced three publications that lie at the intersection of biomedical sciences, computer science, hospital IT, ethics and law, and represent the essence of the interdisciplinary nature of the DPPH project.

The first paper, entitled "Data Protection and Ethics Requirements for Multisite Research with Health Data: A Comparative Examination of Legislative Governance Frameworks and the Role of Data Protection Technologies" was published in the Journal of Law and the Biosciences. This paper compares the legislation on data protection and research ethics requirements for health-related data and personalised medicine across seven jurisdictions. It identifies differences between the European Union and other jurisdictions as a significant barrier for data accessibility in cross jurisdictional multi-site research. Furthermore, it proposes solutions to overcome these legislative differences based on contractual, organizational and technical measures, such as homomorphic encryption and secure multiparty computation.

The second paper, entitled "Cybersecurity of Hospitals: Discussing the Challenges and Working Towards Mitigating the Risks" and published in BMC Medical Informatics and Decision Making, provides a comprehensive description of the main cybersecurity threats related to data management within hospitals. In particular, it recommends addressing cybersecurity through a risk-based approach that manages the tradeoffs between risks and benefits, as well as the different types of risks. Moreover, it discusses the challenges that privacy-conscious data sharing and medical devices pose to security.

Finally, the third paper entitled "Revolutionizing Medical Data Sharing Using Advanced Privacy Enhancing Technologies: Technical, Legal and Ethical Synthesis" was published in the Journal of Medical Internet Research. This paper focused on the use of privacy-enhancing technologies (namely, homomorphic encryption and secure multiparty computation) in medical data sharing from converging technical, ethical and legal perspectives. Its purpose, as part of the project, was to meld scholarship on legal and technological requirements and demonstrate how these two fields can be used to ensure compliance with relevant regulations. In this paper, we also provided a synthesis between two advanced privacy enhancing technologies (PETs): Homomorphic Encryption and Secure Multiparty Computation (defined together as Multiparty Homomorphic Encryption or MHE) and argued that MHE fulfills the legal requirements for medical data sharing under the General Data Protection Regulation (GDPR) which has set a global benchmark for data protection. Specifically, the data processed and shared when using MHE can be considered *anonymized* data. Finally, we explained how MHE can reduce the reliance on customized contractual measures between institutions.

WP2 (Early deployment): This work package prioritized the production of a working prototype (codenamed MedCo) that leverages on the basic building blocks already produced by LDS and DEDIS and could serve as the basis for the subsequent developments in DPPH, aligned with the needs and requirements gathered in WP1. The prototype implemented and integrated during the first year of the project was matured and consolidated in terms of code quality and functionality for cohort discovery and distributed survival analysis in the identified SPHN use cases, with numerous new releases during the reporting period (source code and documentation are available at https://medco.epfl.ch/). This prototype was presented to the University hospitals and to SPHN Data Coordination Center representatives which agreed to launch a test deployment in CHUV, HUG and Inselspital, as foreseen in the companion MedCo project. After that, WP2 continued supporting the refinement and progress on the prototype, considering the feedback received from the MedCo project and initiated the integration of the new primitives developed in WP3.

WP3 (Privacy-Conscious Technology for Medical Data Sharing): This work package focused on the main research activities needed to bring additional functionalities to the basis that the prior work developed at LDS and DEDIS (UnLynx[1]) already provided (counts). This involved multiple courses of action, which were initiated during the first year, and continued during the whole span of the project: (a) research on homomorphic encryption and related cryptographic primitives that enable encrypted processing with limited bandwidth in semi-honest adversarial models (T3.1); (b) research on verification techniques that enable more demanding adversarial models, including input validation (T3.2); (c) research on secret-sharing-based multiparty computation tools to enable efficient statistical computations (T3.3); (d) implementation efforts for the i2b2 database model adaptation (T3.4); and research on novel cryptographic protocols for privacy-preserving data analysis (T3.5).

The work regarding T3.1 and T3.3 focused on the design, implementation and optimization of a lattice-based homomorphic library that supports efficient secure computation and realizes multiparty computation protocols through collectively shared keys, therefore enabling hybrid solutions that leverage the benefits of both homomorphic encryption (low bandwidth, efficient local execution of polynomial functions) and interactive protocols (high versatility, constrained computational complexity). For this purpose, LDS implemented two well-established homomorphic lattice-based cryptosystems (BFV and CKKS) in Golang in a library named Lattigo, and further enhanced them with distributed versions that define efficient interactive protocols for key generation, key switching and decryption in a semi-honest adversarial setting. The library is fully documented and open source (https://github.com/ldsec/lattigo, version 2.3.0 at the time of writing). In 2019, LDS presented the library to the homomorphicencryption.org standardization community and to the audience of the 28th USENIX Security Symposium, and in 2020 demonstrated it to the 8th Workshop on Encrypted Computing and Applied Homomorphic Cryptography (WAHC'20). In both cases, the library was very positively received. The algorithms and protocols for multiparty homomorphic encryption which are implemented in the library are described in two papers (see publications section): (a) a paper entitled "Multiparty Homomorphic Encryption from Ring-Learning-With-Errors" which was published at the Privacy Enhancing Technologies Symposium (PETS'21) and that proposes a solution to the secure-multiparty-computation (MPC) problem using multiparty homomorphic encryption, and (b) a paper entitled "Efficient Bootstrapping for Approximate Homomorphic Encryption with Non-Sparse Keys" which was published at the 40th Annual International Conference on the Theory and Applications of Cryptographic Techniques (EUROCRYPT'21) and which describes a novel efficient bootstrapping procedure for the full-RNS variant of the CKKS cryptosystem.

[1] D. Froelicher, P. Egger, J. S. Sousa, J. L. Raisaro, Z. Huang, C. Mouchet, B. Ford, and J.P. Hubaux, UnLynx: "A Decentralized System for Privacy-Conscious Data Sharing." Privacy Enhancing Technologies Symposium (PETS), volume 4, pages 152–170, Minneapolis, USA, 2017.

The work in T3.2 focused on the design and implementation of a novel solution that enables clients to outsource the data obtained and certified by a data source to a service provider in a privacy- and integrity-preserving manner. To preserve data privacy, our solution relies on the CKKS cryptosystem that supports flexible computations on encrypted data. To protect integrity, it employs lattice-based commitments and zero-knowledge proofs based on the multi-party-computation-in-the-head (or MPC-in-the-head) paradigm; these allow a client to convince the service provider about the correctness of the encrypted data, as well as the authenticity of the underlying plaintext data, using the certification mechanism of the data source. The proposed solution and its experimental evaluation are described in a paper entitled "Privacy and Integrity Preserving Computations with CRISP" that was accepted and presented at the 30th USENIX Security Symposium (Sec'21). Regarding T3.4, the i2b2 database model was modified to accommodate protected medical information and has been integrated to the MedCo platform.

The work done in T3.5 focused on the design and implementation of efficient solutions for machine learning applications on sensitive data, notably in distributed environments. Building on top of the multiparty homomorphic encryption paradigm and the Lattigo library developed in T3.1 and T3.3, we developed solutions that enable training and predictions with generalized linear models and neural networks among N parties, while preserving the data and model confidentiality in a passive adversary model with N-1 collusions. Our solution to the general problem of privacy-preserving distributed learning and its application to generalized linear models is described in a paper entitled "Scalable Privacy-Preserving Distributed Learning" that was accepted and presented at the Privacy Enhancing Technologies Symposium (PETS'21). Accordingly, the solution that accounts for the various cryptographic optimizations required to support neural networks (e.g., multi-layer perceptrons and convolutional neural networks) is described in a paper titled "POSEIDON: Privacy-Preserving Federated Neural Network Learning", which was accepted and presented at the 28th Network and Distributed Systems Symposium (NDSS'21). The latter work was also presented earlier at the Privacy Preserving Machine Learning (PPML'20) Workshop which was (virtually) co-located with the 34th conference on Neural Information Processing Systems (NeurIPS'20). We note that these two works also resulted in two patent applications which have been filed and are currently under examination; they are about to enter the national phase and have been exclusively licensed to Tune Insight (see patents section). Moreover, we demonstrated the seamless application of these solutions to two medical use-cases, namely, Kaplan-Meier survival analysis for oncology and genome-wide association studies for medical genetics. This application resulted in a paper titled "Truly Privacy-Preserving Federated Analytics for Precision Medicine with Multiparty Homomorphic Encryption", which was accepted at Nature Communications. Finally, we showed how the DPPH cryptographic framework can also be used for other genomic tasks and in different settings, i.e., centralized ones. In particular, the solution that was submitted to the iDASH 2019 homomorphic encryption track and involved the execution of a logistic regression model for genotype imputation, is presented in a paper titled "Ultra-Fast Homomorphic Encryption Models enable Secure Outsourcing of Genotype Imputation" which was accepted to the journal Cell Systems.

WP4 (Quantifying and Addressing Inference Risks on Biomedical Databases): This work package targeted the design of an evaluation framework to systematize the analysis of inference attacks that exploit patient data, and produced adequate countermeasures.

LDS, Fellay group and SDSC analyzed potential models to evaluate the various risks stemming from genomic data sharing, and how to quantify them. We produced a comprehensive framework comprising these models, and implemented a proof-of-concept prototype in a decision-support tool that accounts for tangible inference risks, namely kinship inference and phenotype/traits inference. In addition, we carried out a user study to determine the perception and usability of the developed prototype. The results of this work package are described in a paper titled "GenoShare: Supporting Privacy-Informed Decisions for Sharing Individual-Level Genetic Data", which was published at the journal Studies in Health Technology and Informatics and was presented at the Annual Symposium of the American Medical Informatics Association (AMIA) in November 2020.

WP5 (Decentralized Trust and Accountability in Medical Data Processing): This work package emphasized on the development of secure decentralized protocols for access control, logging and accountability purposes. In addition to improving and supporting software libraries such as Cothority, Onet, and Kyber, DEDIS also worked on the design and implementation of a new modular ledger architecture that provides a minimal set of abstractions and module implementations that can be combined to run a distributed ledger. This allows building distributed ledgers that can be easily extended and tested. Furthermore, this work package focused on the design of a fully-decentralized and auditable authorization mechanism as well as a novel architecture for building general-purpose decentralized applications. The authorization mechanism was deployed in MedCo (MedChain), to manage decentralized identity and access control rights in medical settings. The results of these research efforts are summarized in two research publications: a) a paper titled "Rethinking General-Purpose Decentralized Computing" which was accepted and presented at the Workshop on Hot Topics in Operating Systems (HotOS '19), and b) a paper titled "CALYPSO: Private Data Management for Decentralized Ledgers" which was accepted and presented at the 47th International Conference on Very Large Data Bases (VLDB'21).

WP6 (Big Data Infrastructure & Knowledge Management for Medical Data): According to the restructuring of the DPPH work plan towards a project extension, which was filed to and accepted by the PHRT administration (Feb. 17th, 2020), the tasks of WP6 were deprecated and the accompanying budget was split among the other work packages.

WP7 (Patient Monitoring and Data Collection): This work package focused on a prospective analysis of patient-centric tools to access mHealth services in a privacy-conscious way. Its work was finalized early in the project and was consolidated in a paper titled "HideMyApp: Hiding the Presence of Sensitive Apps on Android" which was accepted and presented at the USENIX Security Symposium 2019 (Sec'19). This work comprises an analysis on the information collection carried out by installed apps in the Android ecosystem, as well as a survey and an evaluation of different virtualization techniques for mobile devices that are the core of our privacy-preserving tool for mHealth apps. We built a fully-functional prototype of our tool (see https://hma.epfl.ch/) and evaluated its usability with average users.

WP8 (Deployment and Validation): This work package prioritized the deployment and validation of the privacy preserving tools developed in DPPH, in operational environments. Aligned with the decisions made in WP2, we synchronized this deployment and validation with the SPHN/PHRT-funded project MedCo, receiving the feedback from the three involved university hospitals (CHUV, HUG, and Inselspital) in terms of IT hospital compliance, legal compliance, and usefulness for researchers. In particular, we worked in close collaboration with the SPO (Swiss Personalized Oncology) Driver project of SPHN in which we successfully tested MedCo in 2021. We published a short paper entitled "SPHN/PHRT-MedCo in Action: Empowering the Swiss Molecular Tumor Board with Privacy-Preserving and Real-Time Patient Discovery", in the journal Studies in Health Technology and Informatics to explain how MedCo was used by clinicians of the SPO project. In the Spring of 2021, MedCo was extensively demoed, reviewed and tested by the SPHN National Advisory Board and by the HITSTAG (the IT working group of university hospitals). At the time of this writing, the next step is the launch of 2 pilot projects in 2022, which based on MedCo, will manage data in the SPO and the BioRef projects. Tune Insight will take care of most of the software aspects.

WP9 (Ethics and Users' Perspectives): This package focused on investigating whether new data sharing technologies can offer viable solutions by fostering trustworthy transactions of health-relevant data between researchers, providers, and data subjects. Various layers of this ethical and legal requirements analysis were performed through the course of the DPPH project. First, we conducted a qualitative interview study involving semi-structured interviews with lawyers at stakeholder hospitals and universities, as well as regulatory compliance and clinical data management staff. These interviews involved a set of vignette scenarios involving feasibility study and data erasure requests. For the feasibility study requests, interviewees were required to indicate whether they would accept the request without further approval, deny the request or refer it to an ethics committee. For the data erasure requests, interviewees were asked whether they would completely erase the data or just delete links to the data. Interviewees were also required to rate how important they considered each of these requests on a seven-point Likert scale. This work resulted in a manuscript entitled "Advanced Privacy Enhancing Technologies, Distributed Ledger Technology and Expert Assessments: A Qualitative Interview based Ethico-Legal Study" and which is currently being prepared for submission to the journal *BMC Health Services Research*. Second, we produced a paper that compares patient rights regarding data stored in national electronic health records under nine European and Asia-Pacific jurisdictions. We found that while there is a convergence of access controls, there is a divergence with respect to controlling third-party access and modifying patient data. The paper titled "Whose Health Record? A Comparison of Patient Rights under National Electronic Health Record (NEHR) Regulations in Europe and Asia-Pacific Jurisdictions" was accepted at the Singapore Journal of Legal Studies. Finally, a scoping review was conducted to examine the ethical, legal and socio-technical issues emerging from implementing national eHealth systems. The resulting paper titled "Benefits, Challenges and Contributors to Success for National eHealth Systems Implementation: A Scoping Review" was accepted to the Journal of American Medical Informatics Association (JAMIA). Another paper also examines whether distributed ledger technology such as MedChain can be used to guarantee auditability of patient data. This paper, titled 'Can Distributed Ledger Technology Guarantee the Auditability of Patient Data in Electronic Health Records? A Comparative Legal Study', will be published in the *Journal of Law, Information and Science* in 2022.

Overall achievements

The DPPH project delivered the scientific and technical objectives that were set in its originally envisioned work plan, thus it successfully addressed the privacy, security, and ethical challenges of medical data sharing for effective P4 medicine. In particular, the project delivered a holistic requirements analysis of the medical data sharing ecosystem from the standpoint of legal, ethical, and medical stakeholders as well as developed software-based solutions for accountable and privacy-preserving medical data sharing. Furthermore, it provided tools that enable quantitative analysis of inference risks when sharing patient data as well as a comprehensive ethical analysis of distributed platforms for medical data sharing. In the following, we list the project achievements with respect to its various work packages.

The achievements of WP1 include the completion of the legal and ethical requirements analysis as well as the work on the contextual uses of privacy preserving technologies. In particular, the efforts in WP1 resulted to three cross-disciplinary and complementary papers that highlight: (i) important issues for international multisite medical data sharing that stem from the differences of data protection laws across jurisdictions, (ii) the hospitals' urgent need for addressing cybersecurity threats, and (iii) the prominent role of advanced privacy-enhancing technologies to achieve legal compliance and accelerate medical research.

The achievements of WP2 and WP8 include: (i) the production-ready version of the MedCo software which was evaluated by the SPHN National Advisory Board for potential deployment in the five Swiss University Hospital as part of the Swiss infrastructure for personalized health with a positive outcome, and (ii) the commitment by the SPO SPHN Driver project and the BioRef project to test and validate MedCo as privacy-preserving platform for distributed analytics.

In terms of the privacy-conscious technology developed in WP3, one achievement relates to the publication and evolution of the Lattigo library, that has gained the attention of the international cryptographic community working on homomorphic encryption through its presentation and demonstration at the 8th Workshop on Encrypted Computing and Applied Homomorphic Cryptography (WAHC'20). Moreover, the scientific papers describing the developed solutions for efficient bootstrapping for approximate homomorphic encryption, for input validation in privacy-conscious computations, for privacy-preserving machine learning in distributed environments along with their applications to medical use-cases have been accepted to top-tier information security conferences and biomedical journals. At the system level, a notable achievement is the integration of the produced protection tools in MedCo and its continuous evolution from a privacy-preserving cohort exploration tool to a privacy-preserving medical analysis platform.

WP4 and WP7 produced tools that can be directly used by patients to (a) make informed decisions and grant or deny consent to access their medical data on the bases of quantitative and visually representative measures of risk, and to (b) avoid undesired privacy leakage from health-related applications to other third-party applications installed on their connected devices (such as smartphones and tablets). The papers describing these tools have also been accepted at a medical journal and a top-tier information security conference, respectively.

Scientific achievements for WP5 include the design and implementation of an auditable private data sharing and management platform on decentralized ledgers. The paper describing this solution was accepted at a top-tier database conference. This system was also adapted for enabling its use in conjunction with MedCo (MedChain), to manage decentralized identity and access control rights in medical settings. Finally, scientific highlights for WP9 include the completion of the qualitative data collection, the presentation of a conference paper on comparative national electronic health record regulations and its subsequent acceptance to the journal of American Medical Informatics Association, as well as the acceptance of a paper on comparative auditability requirements to the Singapore Journal of Legal Studies.

How do your rate the achievement of your project:
underachieved / achieved as expected / overachieved? Please explain.

As discussed in the above section, the DPPH project successfully achieved the scientific and technical objectives that were specified in its original work plan with a series of technical developments and top-tier scientific publications. Due to external events impacting the project execution, however, the DPPH project did not build a scalable scientific infrastructure for medical data on top of the Swiss Data Science Center (SDSC) computing platform that supports big data processing and knowledge management. The project PIs in collaboration with the PHRT administration agreed on a modified project proposal that reduced the contributions of WP6 and redistributed the budget to other WPs, leading to a zero-cost project extension. Overall, taking into account the modified project proposal, we rate the DPPH project as overachieved. In particular, the launch of a start-up company (Tune insight) based on the outcomes of the project goes beyond what was initially envisioned.

What kind of help / support would have made the collaboration in your project easier?

The DPPH consortium valued the efforts made by the PHRT to enable collaboration, and found them appropriate and fit. In particular, we consider that DPPH did certainly benefit from PHRT's contribution to the awareness of data protection and privacy challenges within PHRT/SPHN projects. Yet, what was missing, and which we reported to the PHRT administration, was the obligation to each project to set up and maintain a decent website. This is an indispensable tool to foster collaboration within the program.

Here is some evidence:

https://www.sfa-phrt.ch/approved-projects-2017

In spite of these projects being 4 years old, only 2 (PSSS and DPPH) have a displayed website. The time and energy spent in the writing of reports such as this one (that are accessible to and read by very few people) would be much better invested by setting up websites (that are accessible to everyone).

Was your financing sufficient? If not, how much of your financing was lacking and was obtained possibly from other sources? 5%, 10%... ? Please specify.

The financing provided to the DPPH consortium was sufficient and appropriate. In addition, we acknowledge the cooperation and flexibility of the PHRT administration to accept a modified version of the original project proposal (that reduced the contributions of WP6 to adapt to external events impacting the project execution) and which led to a redistribution of the project funds and to a project extension (Feb. 17th, 2020). Also, it turned out that the

Ethics group of ETH led by Effy Vayena managed to wrap up their contributions faster than planned. Consequently, a fund reallocation was granted on December 1st, 2020.

If you would redo the project, what would you do differently? Please explain.

As discussed above, the DPPH project successfully achieved its primary objectives. While external events delayed one of its envisioned goals, in particular the development of a scalable scientific computing infrastructure on top of the SDSC framework that enables data traceability and knowledge management (WP6), these did not affect the project's overall progress. The consortium's decision along with the cooperation of the PHRT administration to reduce the WP6 contributions was reasonable and correct, and allowed the project to reach its end smoothly. As such, we would not do things differently but only wish that these external events had not existed.

Outlook

Activities related to the project will continue in the following way:

On the research front, the EPFL/LDS laboratory will continue its efforts on including additional functionalities to the framework for secure and privacy-conscious medical data sharing for analytics. In particular, we will expand the framework with support for machine learning preprocessing pipelines, such as hyperparameter tuning and principal component analyses, in federated settings. Moreover, we will extend the framework's algorithmic toolkit with state-of-the-art machine learning models such as deep convolutional neural networks and recurrent ones. Our objective is to leverage on these new features to enable more complex medical use-cases such as genome wide association studies (GWAS) and single-cell analyses. Furthermore, we are currently working on verifiable computation techniques that would allow the DPPH framework to operate under strong threat models, e.g., where some of the involved parties are behaving maliciously. Finally, we will disseminate the Lattigo cryptographic library with the publication of a whitepaper.

As already mentioned, Tune Insight will further develop the product and will provide support for deployment and software maintenance, in health-related applications and beyond.

ETH/Vayena Lab will continue studying the ethical implications of personalized medicine.

The EPFL/DEDIS laboratory will continue the development of secure and privacy-preserving decentralized protocols towards a long-term effort to build a personhood-based decentralized digital platform to enable human participation in self-governing communities. In addition, we will maintain and further develop Calypso. In particular, two possibilities of extending it are: (a) so that the data can not only be stored but also "computed on" via MPC for example, and (b) so that the secret-management policies themselves can be private (and not just identity-anonymized) and MPC might be used for evaluating whether a particular usage of data is actually authorized or not. A use-case for "private Calypso policies" might fit into our coming research project PAIDIT with the International Committee of the Red Cross: e.g., policies for the management and key recovery for wallets in ICRC's custody.

EPFL/Fellay Lab will keep promoting the DPPH tools and their application to concrete use cases as well as collaborating with EPFL/LDS and Tune Insight for further developing

MedCo by providing new use cases and functional requirements. There exist ongoing discussions with international partners in the Netherlands, Italy and USA for using MedCo in the framework of multicentric projects in precision medicine research.

Networking and collaboration with other PHRT-, SPHN- and/or SDSC-entities

The DPPH project approached other SPHN/PHRT projects, including SVIP-O, PRECISE, PSSS, SPO, and TIMES, in order to present the privacy solutions produced and to look for further application scenarios and particular use cases that can make use of the DPPH platform with the aim to enhance the security and privacy of data sharing in SPHN. The reception of DPPH, especially of the MedCo tool, was very positive, and these conversations have revealed promising synergies and collaboration opportunities. In order to optimize the resources available in DPPH, the consortium decided to focus their efforts on the most promising use cases, the already mentioned SPO and BioRef projects.

Another notable collaboration was the Secure Collective Covid-19 Research initiative, a consortium set up to support international collaborations on COVID-19 research while respecting patient privacy using the MedCo platform. A whitepaper describing the efforts on this front has been published at the Journal of the American Medical Informatics Association. Finally, the DPPH project has also been in contact with other European initiatives and projects focusing on similar goals, e.g., the Personal Health Train, for exchanging ideas and opinions.

Data Management: how did you process, store, exchange and archive the project data?

The DPPH consortium established a shared SWITCHdrive folder where DPPH researchers could access all the data needed for the project's correct development. The meeting agendas, the minutes and the used project slides were kept in a Google Drive folder shared only within the consortium. Additionally, for any data collected within the questionnaires carried out in WP1, the project maintained a specific folder, secured via access permissions so that only the directly related project staff could access it. For the testing and evaluation phase, the project had specific and temporary data stores in-premise, used by each of the groups performing the tests. All of the above data has been deleted since the tests are now finalized.

For the experimental evaluation sections of the papers describing the privacy-conscious technologies (WP3-WP5), the project employed publicly-available datasets that can be found online in various data repositories.

All data for the stakeholder interviews in WP9 were stored on a pair of password-protected USB drives that are held at the Health Ethics and Policy Laboratory offices in Zurich. The transcribed interviews were stored on these USB drives, as well as on the internal network storage drives for staff at ETH. However, for additional security purposes, and in compliance with the project's ethics requirements, all of these interviews were pseudonymised. All interviewee identifiers in the interview transcript have been replaced with a randomly generated code. Any other identifiers have been removed and replaced with a blank value.

Part B - Items that will be consolidated from all individual PHRT projects each year

List involved medical institutions

Name	Address	Place
CHUV (Medco Project)	Rue du Bugnon 21, CH-1011	Lausanne, VD
Inselspital (Medco Project)	Freiburgstrasse, CH-3010	Bern, BE
HUG (Medco Project)	Rue Gabrielle-Perret-Gentil 4, CH-1205	Genève, GE

List the three most important contributions of the project to strengthen the collaboration with the medical-clinical sector

A very significant first contribution of the project was the legal and ethics requirements analysis for multi-site research involving health data, which was produced as a result of the work done in WP1 and WP9. This analysis is of extreme importance for future research initiatives aiming to strengthen their medical results via data sharing among multiple medical institutions. A second important contribution is the early prototype produced in WP2, which enabled the medical-clinical sector to test and evaluate privacy-preserving mechanisms in the ongoing PHRT/SPHN technology transfer project titled MedCo. Finally, a third contribution is the DPPH framework that includes solutions for privacy conscious machine learning, inference analysis, and secure distributed access control, developed in WP3, WP4, and WP5; these will enable expanding the functionalities and improving the versatility of the MedCo prototype, as well as adapt it to the needs and requirements of the Swiss medical-clinical sector.

How many patients incl. responsible physician & hospital participated / contributed to this project?

Not applicable.

Part C - Annex

Publication list out of the PHRT project

i) All publications acknowledging PHRT explicitly

A. Pham, I. Dacosta, E. Losiouk, J. Stephan, K. Huguenin, and J-P. Hubaux, HideMyApp: Hiding the Presence of Sensitive Apps on Android, in 28th USENIX Security Symposium, Santa Clara, California, USA, 2019

JL. Raisaro, JR. Troncoso-Pastoriza, S. Pradervand, M. Cuendet, M. Misbach, J. Sa, F. Marino, N. Freundler, N. Rosat, D. Cavin, A. Leichtle, J. Fellay, O. Michielin, and J-P. Hubaux, SPHN/PHRT-MedCo in Action: Empowering the Swiss Molecular Tumor Board with Privacy-Preserving and Real-Time Patient Discovery, Studies in Health Technology and Informatics, 2020 Jun 1;270:1161-2

S. Sav, A. Pyrgelis, JR. Troncoso-Pastoriza, D. Froelicher, J-P. Bossuat, J. Sa Sousa, and J-P. Hubaux, POSEIDON: Privacy-Preserving Federated Neural Network Learning, in

Privacy-Preserving Machine Learning Workshop (PPML), co-located with the 34th Conference on Neural Information Processing Systems (NeurIPS), 2020

JL. Raisaro, JR. Troncoso-Pastoriza, S. Pradervand, M. Cuendet, M. Misbach, J. Sa, F. Marino, N. Freundler, N. Rosat, D. Cavin, A. Leichtle, J. Fellay, O. Michielin, and J-P. Hubaux, SPHN/PHRT-MedCo in Action: Empowering the Swiss Molecular Tumor Board with Privacy-Preserving and Real-Time Patient Discovery, Studies in Health Technology and Informatics, 2020 Jun 1;270:1161-2

J. Scheibner, M. Ienca, S. Kechagia, JR. Troncoso-Pastoriza, JL. Raisaro, J-P. Hubaux, and E. Vayena, Data Protection and Ethics Requirements for Multisite Research with Health Data: A Comparative Examination of Legislative Governance Frameworks and the Role of Data Protection Technologies, Journal of Law and the Biosciences, 2020

C. Mouchet, J-P. Bossuat, JR. Troncoso-Pastoriza, and J-P. Hubaux, Lattigo: a Multiparty Homomorphic Encryption Library in Go, in 8th Workshop on Encrypted Computing and Applied Homomorphic Cryptography (WAHC), 2020

S. Carpov, N. Gama, M. Georgieva, and JR. Troncoso-Pastoriza, Privacy-preserving Semi-parallel Logistic Regression Training with Fully Homomorphic Encryption, BMC Medical Genomics 13, 88, 2020

S. Sav, A. Pyrgelis, JR. Troncoso-Pastoriza, D. Froelicher, J-P. Bossuat, J. Sa Sousa, and J-P. Hubaux, POSEIDON: Privacy-Preserving Federated Neural Network Learning, in Proceedings of the 28th Networks and Distributed Systems Security Symposium (NDSS), 2021

D. Grishin, JL. Raisaro, JR. Troncoso-Pastoriza, K. Obbad, K. Quinn, M. Misbach, J. Gollhardt, J. Sa Sousa, J. Fellay, GM. Church, and J-P. Hubaux, Citizen-centered, Auditable and Privacy-preserving Population Genomics, Nature Computational Science, 2021, Mar;1(3):192-8

J. Scheibner, M. Ienca, and E. Vayena, Whose health record?: A Comparison of Patient Rights under National Electronic Health Record (NEHR) Regulations in Europe and Asia-pacific Jurisdictions, Singapore Journal of Legal Studies, Apr 2021:56-75

M. Kim, A. Harmanci, J-P. Bossuat, S. Carpov, J. H. Cheon, I. Chillotti, W. Cho, D. Froelicher, N. Gama, M. Georgieva, S. Hong, J-P. Hubaux, D. Kim, K. Lauter, Y. Ma, L. Ohno-Machado, H. Sofia, Y. Son, Y. Song, JR. Troncoso-Pastoriza, and X. Jiang, Ultra-Fast Homomorphic Encryption Models enable Secure Outsourcing of Genotype Imputation, Cell Systems, 2021

C. Mouchet, JR. Troncoso-Pastoriza, J-P. Bossuat, and J-P. Hubaux, Multiparty Homomorphic Encryption from Ring-Learning-With-Errors, in Proceedings on Privacy Enhancing Technologies (PETS), 2021

D. Froelicher, JR. Troncoso-Pastoriza, A. Pyrgelis, S. Sav, J. Sa Sousa, J-P. Bossuat, and J-P. Hubaux, Scalable Privacy-Preserving Distributed Learning, in Proceedings on Privacy Enhancing Technologies (PETS), 2021

S. Chatel, A. Pyrgelis, JR. Troncoso-Pastoriza, and J-P. Hubaux, Privacy and Integrity Preserving Computations with CRISP, in 30th USENIX Security Symposium, Vancouver, B.C., Canada, 2021

E. Kokoris-Kogias, E. Ceyhun Alp, L. Gasser, P. Jovanovic, E. Syta, and B. Ford, CALYPSO: Private Data Management for Decentralized Ledgers, in Proceedings of the 47th International Conference on Very Large Data Bases (VLDB), 2021

J. Scheibner, J. Sleigh, M. Ienca, and E. Vayena, Benefits, Challenges, and Contributors to Success for National eHealth Systems Implementation: A Scoping Review, Journal of the American Medical Informatics Association, Volume 28, Issue 9, September 2021, Pages 2039–2049

A. Senf, R. Davies, F. Haziza, J. Marshall, JR. Troncoso-Pastoriza, O. Hofmann, T. M. Keane, Crypt4GH: a File Format Standard Enabling Native Access to Encrypted Data, Bioinformatics, Volume 37, Issue 17, 1 September 2021, Pages 2753–2754

J-P. Bossuat, C. Mouchet, JR. Troncoso-Pastoriza, and J-P. Hubaux, Efficient Bootstrapping for Approximate Homomorphic Encryption with Non-Sparse Keys, in 40th Annual International Conference on the Theory and Applications of Cryptographic Techniques (EUROCRYPT), 2021

D. Froelicher, JR. Troncoso-Pastoriza, JL. Raisaro, M. Cuendet, J. Sa Sousa, H. Cho, B. Berger, J. Fellay, and J-P. Hubaux, Truly Privacy-Preserving Federated Analytics for Precision Medicine with Multiparty Homomorphic Encryption, Nature Communications 12, 5910, 2021

ii) Publications of at least two groups of the PHRT project and joint papers with other PHRT, SPHN or SDSC projects

D. Froelicher, M. Misbach, JR. Troncoso-Pastoriza, JL. Raisaro, and J-P. Hubaux, MedCo2: Privacy-Preserving Cohort Exploration and Analysis, Studies in Health Technology and Informatics, 2020 Jun 1;270:317-21

JL. Raisaro, JR. Troncoso-Pastoriza, Y. El-Zein, M. Humbert, C. Troncoso, J. Fellay, and J-P. Hubaux, GenoShare: Supporting Privacy-Informed Decisions for Sharing Individual-Level Genetic Data, Studies in Health Technology and Informatics, 2020, Jun 1;270:238-41

J. Scheibner, M. Ienca, S. Kechagia, JR. Troncoso-Pastoriza, JL. Raisaro, J-P. Hubaux, and E. Vayena, Data Protection and Ethics Requirements for Multisite Research with Health Data: A Comparative Examination of Legislative Governance Frameworks and the Role of Data Protection Technologies, Journal of Law and the Biosciences, 2020

JL. Raisaro, JR. Troncoso-Pastoriza, S. Pradervand, M. Cuendet, M. Misbach, J. Sa, F. Marino, N. Freundler, N. Rosat, D. Cavin, A. Leichtle, J. Fellay, O. Michielin, and J-P. Hubaux, SPHN/PHRT-MedCo in Action: Empowering the Swiss Molecular Tumor Board with Privacy-Preserving and Real-Time Patient Discovery, Studies in Health Technology and Informatics, 2020 Jun 1;270:1161-2

J. Scheibner, JL. Raisaro, JR. Troncoso-Pastoriza, M. Ienca, J. Fellay, E. Vayena, and J-P. Hubaux, Revolutionizing Medical Data Sharing Using Advanced Privacy Enhancing Technologies: Technical, Legal and Ethical Synthesis, Journal of Medical Internet Research, 2021

D. Froelicher, JR. Troncoso-Pastoriza, JL. Raisaro, M. Cuendet, J. Sa Sousa, H. Cho, B. Berger, J. Fellay, and J-P. Hubaux, Truly Privacy-Preserving Federated Analytics for Precision Medicine with Multiparty Homomorphic Encryption, Nature Communications 12, 5910, 2021

iii) Other publications with minor contributions of the PHRT project

E. Ceyhun Alp, E. Kokoris-Kogias, G. Fragkouli, and B. Ford, Rethinking General-Purpose Decentralized Computing, in Proceedings of the Workshop on Hot Topics in Operating Systems (HotOS '19), Association for Computing Machinery, New York, NY, USA, 105–112, 2019

C. Boura, N. Gama, M. Georgieva, and D. Jetchev, Simulating Homomorphic Evaluation of Deep Learning Predictions, in International Symposium on Cyber Security Cryptography and Machine Learning, Springer, Cham, 2019

C. Boura, N. Gama, M. Georgieva, and D. Jetchev, Chimera: Combining Ring-lwe-based Fully Homomorphic Encryption Schemes, Journal of Mathematical Cryptology, Jan 1;14(1):316-38, 2020

I. Chillotti, N. Gama, M. Georgieva, and M. Izabachène, TFHE: Fast Fully Homomorphic Encryption Over the Torus, Journal of Cryptology, Jan;33(1):34-91, 2020

M. Ienca and E. Vayena, On the responsible use of digital data to tackle the COVID-19 pandemic. Nature Medicine, 26(4), 463-464, 2020

U. Gasser, M. Ienca, J. Scheibner, J. Sleigh, and E. Vayena, Digital tools against COVID-19: Taxonomy, Ethical Challenges, and Navigation Aid, The Lancet Digital Health, 2020

M. Ienca and J. Scheibner, What is Neurohacking? Defining the Conceptual, Ethical and Legal Boundaries, Ethical Dimensions of Commercial and DIY Neurotechnologies, 3, 203, 2020

A. Pedrouzo-Ulloa, JR. Troncoso-Pastoriza, N. Gama, M. Georgieva, and F. Pérez-González, Multiquadratic Rings and Walsh-Hadamard Transforms for Oblivious Linear Function Evaluation, in IEEE International Workshop on Information Forensics and Security (WIFS), 2020, pp. 1-6

J. Scheibner, A. Jobin, and E. Vayena, Ethical Issues with Using Internet of Things Devices in Citizen Science Research: A Scoping Review, Cambridge Handbook of Life Science, Information Technology and Human Rights, 2020

RD. Gosselin, C. Redin, É. Ristorcelli, JL. Raisaro, and J. Fellay, Precision Medicine: at the Crossroad of Molecular Biology, Information Sciences and Humanities, Revue Medicale Suisse, 2020 Sep 1;16(704):1574-8

C. Baum, D. Escudero, A. Pedrouzo-Ulloa, P. Scholl P, and JR. Troncoso-Pastoriza, Efficient Protocols for Oblivious Linear Function Evaluation from Ring-LWE, In International Conference on Security and Cryptography for Networks 2020 Sep 14 (pp. 130-149). Springer, Cham.

ST. Argaw, JR. Troncoso-Pastoriza, D. Lacey, MV. Florin, F. Calcavecchia, D. Anderson, W. Burleson, JM. Vogel, C. O'Leary, B. Eshaya-Chauvin, and A. Flahault, Cybersecurity of Hospitals: Discussing the Challenges and Working Towards Mitigating the Risks, BMC Medical Informatics and Decision Making, 2020 Dec;20(1):1-0

D. Froelicher, JR. Troncoso-Pastoriza, JS Sousa, and J-P. Hubaux, Drynx: Decentralized, Secure, Verifiable System for Statistical Queries and Machine Learning on Distributed Datasets, in IEEE Transactions on Information Forensics and Security, 2020 Mar 2;15:3035-50

JL Raisaro, F. Marino, JR. Troncoso-Pastoriza, R. Beau-Lejdstrom, R. Bellazzi, R. Murphy, E. Bernstam, H. Wang, M. Bucalo, Y. Chen, A. Gottlieb, A. Harmanci, M. Kim, Y. Kim, J. Klann, C. Klersy, B. Malin, M. Méan, F. Prasser, L. Scudeller, A. Torkamani, J. Vaucher, M. Puppala, S. Wong, M. Frenkel-Morgenstern, H. Xu, B. Maiyaki Musa, A. Habib, T. Cohen, A. Wilcox, H. Salihu, H. Sofia, X. Jiang, and J-P. Hubaux, SCOR: A Secure International Informatics Infrastructure to Investigate COVID-19, in Journal of the American Medical Informatics Association, Volume 27, Issue 11, November 2020, Pages 1721–1726

C. Baum, D. Escudero, A. Pedrouzo-Ulloa, P. Scholl P, and JR. Troncoso-Pastoriza, Efficient Protocols for Oblivious Linear Function Evaluation from Ring-LWE, Journal of Computer Security, IOS Press, January, 2021

A. Pedrouzo-Ulloa, JR. Troncoso-Pastoriza, N. Gama, M. Georgieva, and F. Pérez-González, Revisiting Multivariate Ring Learning with Errors and Its Applications on Lattice-Based Cryptography, Mathematics, 2021; 9(8):858

S. Chatel, A. Pyrgelis, JR. Troncoso-Pastoriza, and J-P. Hubaux, SoK: Privacy-Preserving Collaborative Tree-based Model Learning, in Proceedings on Privacy Enhancing Technologies (PETS), 2021

Media releases, publications/articles in the public media (if any)

https://www.letemps.ch/economie/lepfl-lance-un-logiciel-exploiter-maniere-sure-donnees-medicales

https://actu.epfl.ch/news/a-cryptography-game-changer-for-biomedical-researc/

Preliminary developments, ideas or insights that might lead to applications, new projects or patents

Most of the developments and ideas that correspond to the solutions, tools and technologies developed in WP3, WP4, and WP5, are mature enough and will be consolidated through a startup company (see next section).

Licenses, patents and spin-offs (if any)

B. Ford, L. Gasser, E Kokoris-Kogias, and P. Jovanovic, PCT/EP2018/053886 “Methods and Systems for Secure Data Exchange”, 2018

D. Froelicher, JR. Troncoso-Pastoriza, A. Pyrgelis, S. Sav, J. Sa Sousa, J-P. Bossuat, and J-P. Hubaux, PCT/EP2020/062810 “System and Method for Privacy-Preserving Distributed Training of Machine Learning Models on Distributed Datasets”, 2020

S. Sav, JR. Troncoso-Pastoriza, A. Pyrgelis, D. Froelicher, J-P. Bossuat, J. Sa Sousa, and J-P. Hubaux, PCT/EP2020/074031 “System and Method for Privacy-Preserving Distributed Training of Neural Network Models on Distributed Datasets”, 2020

B. Ford, US2021018953 (A1), “Asynchronous Distributed Coordination and Consensus with Threshold Logical Clocks”, 2021

Tune Insight SA, a startup incubated at the EPFL Laboratory for Data Security (LDS), https://tuneinsight.com/

Others / comments

Not applicable.