Deutsche Version (siehe unten)
Version française (ci-dessous)

Lay Summary

Project title	Data Protection in Personalized Health (DPPH)
Main applicant	Prof. Jean-Pierre Hubaux, EPFL
Consortium	Prof. Jean-Pierre Hubaux (EPFL), Prof. Bryan Ford (EPFL), Prof. Dimitar Jetchev (EPFL), Prof. Jacques Fellay (EPFL), Prof. Effy Vayena (ETHZ), Dr. Olivier Verscheure (SDSC)
Short Summary	DPPH addresses 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. The main result of the project will be a platform composed of software packages that seamlessly enable clinical and genomic data sharing and exploitation across a federation of medical institutions, hospitals and research laboratories across Switzerland in a scalable, secure, responsible and privacy-conscious way, and that integrates widespread cohort exploration tools and analysis frameworks.
Background	P4 (Predictive, Preventive, Personalized and Participatory) medicine is called to revolutionize healthcare by providing better diagnoses and targeted preventive and therapeutic measures. However, to accelerate its adoption and maximize its potential, clinical and research data on large numbers of individuals must be efficiently shared between all stakeholders. The privacy risks stemming from disclosing medical data raise serious concerns, and have become a barrier that can hold back the advances in P4 medicine if effective privacy-preserving technologies are not adopted to enable privacy-conscious medical data sharing. The evolution of the regulation towards further guarantees (e.g., HIPAA in USA and the new GDPR in EU) reflects this urgent need. The combination of data sharing with recent advances in the field of *omics and, in particular, in high-throughput sequencing technology, leads to an explosive growth in the amounts of available data; this big data scale can usually not be handled with current hospital computing facilities, hence the need for elastic computing resources that can cope with huge amounts of data in a secure and privacy-aware infrastructure, supporting data processing and sharing.
Goal	DPPH seeks to address 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 deploying an appropriate set of computing tools to make it happen. This main goal materializes in the following outcomes that the project expects to deliver: (i) A holistic requirements analysis of the medical data sharing ecosystem, from the standpoint of legal, ethical and medical stakeholders, (ii) a scalable scientific computing infrastructure, building on top of Swiss Data Science Center's (SDSC) data science framework, (iii) software-based solutions for accountable

	and privacy-preserving data sharing featuring trust distribution across a federation of sites with no single points of failure, (iv) a quantitative analysis of inference risks, and countermeasures for addressing them when releasing aggregated results on patient data, and (v) a comprehensive ethical analysis of distributed platforms for medical data sharing.
Significance	DPPH is meant to combine knowledge from the data science, computer science, ethics, medicine and genomics communities to effectively tackle the challenges currently thwarting data sharing for P4 medicine. The software platforms and prototypes produced by the project are meant to be enablers that effectively combine secure and privacy-conscious data access and processing with large-scale collaborative medical research, addressing the main technological barriers holding up advances personalized medicine. The privacy and ethical frameworks enable an in-depth analysis and evaluation of current and future systems, allowing for future-proofness of the used platforms under current and upcoming strict regulatory frameworks. By establishing liaisons with other PHRT/SPHN projects, DPPH seeks to cover the Swiss national level, targeting prototypes at a national scale, and by leveraging on already established connections with the Global Alliance for Genomics and Health (GA4GH) and its Software Security Group, DPPH also guarantees international relevance and consistency.

Deutsch

Projekttitel	Data Protection in Personalized Health (DPPH)
Hauptgesuchssteller	Prof. Jean-Pierre Hubaux, EPFL
Consortium	Prof. Jean-Pierre Hubaux (EPFL), Prof. Bryan Ford (EPFL), Prof. Dimitar Jetchev (EPFL), Prof. Jacques Fellay (EPFL), Prof. Effy Vayena (ETHZ), Dr. Olivier Verscheure (SDSC)
Kurzzusammen-fassung	DPPH befasst sich mit den wichtigsten Herausforderungen des Datenschutzes, der Sicherheit, der Skalierbarkeit und der Ethik des Datenaustauschs. Ziel ist es, eine effektive P4-Medizin zu ermöglichen, indem ein optimales Gleichgewicht zwischen Benutzerfreundlichkeit, Skalierbarkeit und Datenschutz definiert wird. Das Hauptergebnis des Projekts wird eine Softwareplattform sein, die einen nahtlosen, zuverlässigen und datenschutzfreundlichen Austausch und eine entsprechende Nutzung von klinischen und genomischen Daten durch einen Zusammenschluss von medizinischen Einrichtungen, Spitälern und Forschungslabors in der ganzen Schweiz hinweg ermöglicht und ferner verbreitete Kohortenexplorations und -analysewerkzeuge integriert.
Hintergrund	Man geht davon aus, dass die P4-Medizin («Predictive», «Preventive», «Personalized» und «Participatory») die medizinische Versorgung durch bessere Diagnosen und gezielte präventive und therapeutische Massnahmen revolutionieren wird. Um die Akzeptanz zu beschleunigen und das Potenzial zu maximieren, ist es jedoch unabdingbar, dass grosse Mengen an klinischen und Forschungsdaten effizient zwischen allen Beteiligten ausgetauscht werden. Es versteht sich von selbst, dass dies grosse Bedenken hinsichtlich des Datenschutzes mit sich zieht. Ohne wirksame Technologien für den sicheren Austausch medizinischer Daten können diese Bedenken zum Stolperstein für die Forschung werden und die Fortschritte in der P4-Medizin bremsen könnten. Die Entwicklung von Vorschriften zu mehr Schutz (z.B. HIPAA in den USA und das neue GDPR in der EU) spiegelt die entsprechende dringliche Notwendigkeit wieder. Der steigende Datenaustausch und die jüngsten Fortschritte im –omics-Bereich, insbesondere in den High-Throughput-Sequenzierungstechnologien, führt überdies zu einem explosionsartigen Wachstum an verfügbaren Datenmengen; diese grosse Datenmenge kann in der Regel nicht von der vorhandenen IT-Infrastruktur in Spitälern bewältigt werden. Aus diesem Grund sind anpassungsfähige IT-Ressourcen essentiell, um entsprechende Datenmengen nicht nur zu stemmen, sondern auch deren Austausch und Weiterverarbeitung auf sichere und datenschutzgerechte Weise zu ermöglichen.
Ziel	Das DPPH-Projekt zielt darauf ab, die wichtigsten Herausforderungen des Datenschutzes, der Sicherheit, der Skalierbarkeit und der Ethik des Datenaustauschs für eine effektive P4-Medizin anzugehen, indem ein optimales Gleichgewicht zwischen Benutzerfreundlichkeit, Skalierbarkeit

	und Datenschutz definiert und ein geeignetes Set von Computerwerkzeugen eingesetzt wird. Dieses Hauptziel manifestiert sich in folgenden Projektzielen: (i) eine ganzheitliche Analyse der Bedürfnisse des Ökosystems für den Austausch von Gesundheitsdaten aus der Sicht der rechtlichen, ethischen und medizinischen Interessengruppen, (ii) eine erweiterbare wissenschaftliche IT-Infrastruktur, welche die Softwareplattform des Swiss Data Science Center (SDSC) ergänzt, (iii) fehlerfreie softwarebasierte Lösungen für einen verantwortungsbewussten und datenschutzkonformen Datenaustausch innerhalb der Interessensgruppen, (iv) eine quantitative Analyse der Inferenzrisiken, die bei der Veröffentlichung von aggregierten Resultaten von Patientendaten auftreten können, und entsprechende Gegenmassnahmen sowie (v) eine umfassende ethische Analyse von Distributionsplattformen für den Austausch medizinischer Daten.
Bedeutung	Das DPPH-Projekt hat zum Ziel, das Wissen aus den Bereichen Datenwissenschaft, Informatik, Ethik, Medizin und Genomik zu vereinen und so die Herausforderungen, die den Datenaustausch für die P4-Medizin derzeit behindern, effektiv anzugehen. Die im Rahmen des Projekts entwickelten Softwareplattformen und Prototypen sollen den Zugriff und die Weiterverarbeitung von Daten für grossangelegte medizinische Forschungskooperationen sicher und datenschutzgerecht ermöglichen und die wichtigsten technologischen Barrieren für die Weiterentwicklung der personalisierten Medizin beseitigen. Die Rahmenbedingungen hinsichtlich Datenschutz und Ethik erlauben eine eingehende Analyse und Evaluation derzeitiger und zukünftiger Systeme und gewährleisten so die Nachhaltigkeit der Plattformen unter Berücksichtigung der aktuellen und zukünftigen rechtlichen Rahmenbedingungen. Durch die Verknüpfung mit anderen PHRT/SPHN-Projekten strebt DPPH eine Abdeckung aller Bedürfnisse in der Schweiz an, angefangen bei Prototypen auf nationaler Ebene. Das DPPH-Projekt stellt auch die internationale Relevanz und Kohärenz sicher und baut auf bestehenden Verbindungen zur Global Alliance for Genomics and Health (GA4GH) und deren IT-Sicherheitsgruppe auf.

Français

Titre du projet	Data Protection in Personalized Health (DPPH)
Requérant principal	Prof. Jean-Pierre Hubaux, EPFL
Consortium	Prof. Jean-Pierre Hubaux (EPFL), Prof. Bryan Ford (EPFL), Prof. Dimitar Jetchev (EPFL), Prof. Jacques Fellay (EPFL), Prof. Effy Vayena (ETHZ), Dr. Olivier Verscheure (SDSC)
Résumé	Le projet DPPH s'adresse aux problèmes principaux de confidentialité, de sécurité, d'extensibilité et d'éthique liés au partage des données pour permettre une médecine P4 efficace, en définissant un équilibre optimal entre simplicité, extensibilité et protection des données. Le résultat principal du projet sera une plateforme composée de logiciels permettant un partage et une exploitation de données cliniques et génomiques de manière sécurisée, fiable, et respectueuse de la vie privée, à travers une fédération d'établissements médicaux, d'hôpitaux et de laboratoires de recherche dans toute la Suisse, et qui intègre des outils d'exploration de cohortes ainsi que des outils d'analyse.
Contexte	La médecine P4 (Prédictive, Préventive, Personnalisée et Participative) est appelée à révolutionner les soins médicaux en fournissant de meilleurs diagnostics et des mesures préventives et thérapeutiques ciblées. Cependant, pour accélérer son adoption et maximiser son potentiel, les données cliniques et de recherche sur un grand nombre d'individus doivent être efficacement partagées entre tous les participants. Les risques d'atteinte à la vie privée découlant de la divulgation de données médicales suscitent de graves préoccupations et sont devenus un obstacle pouvant freiner les progrès de la médecine P4 si des technologies efficaces de protection de la vie privée ne sont pas adoptées. L'évolution de la réglementation vers plus de protection (par exemple, HIPAA aux États-Unis et le nouveau RGPD dans l'UE) reflète ce besoin urgent. La combinaison du partage des données avec les progrès récents dans le domaine des -omiques et, en particulier, dans la technologie de séquençage à haut débit, entraîne une croissance exponentielle des quantités de données disponibles; cette grande échelle de données ne peut généralement pas être gérée avec les installations informatiques hospitalières actuelles, d'où la nécessité de ressources informatiques adaptables, capables de gérer d'énormes quantités de données dans une infrastructure sécurisée et confidentielle, prenant en charge le traitement et le partage des données.
But	Le projet DPPH cherche à résoudre les principaux problèmes de partage de données liés à la confidentialité, à la sécurité, à l'extensibilité et à l'éthique en définissant un équilibre optimal entre facilité d'utilisation, performance, et protection des données et en déployant un ensemble d'outils informatiques adapté. Ce projet prévoit d'accomplir les points suivants: (i) Une analyse holistique des besoins de l'écosystème du partage des données médicales, du point de vue des acteurs

	juridiques, éthiques et médicaux, (ii) une infrastructure informatique scientifique extensible, qui complémente la plateforme logicielle proposée par le Swiss Data Science Center (SDSC), (iii) des solutions logicielles de partage de données respectueuses de la vie privée en utilisant une fédération de sites sans point de faute unique; (iv) une analyse quantitative des risques d'inférence lors de la publication de résultats agrégés sur des données de patients, ainsi que des contre-mesures pour y remédier, et (v) une analyse éthique complète des plateformes distribuées pour le partage de données médicales.
Importance	Le projet DPPH vise à combiner les connaissances issues des communautés de la *data science*, de l'informatique, de l'éthique, de la médecine et de la génomique pour relever efficacement les défis qui entravent actuellement le partage de données pour la médecine P4. Les plateformes logicielles et les prototypes produits par le projet se veulent des facilitateurs qui combinent efficacement l'accès et le traitement de données sécurisés et confidentielles pour la recherche médicale collaborative à grande échelle, en s'attaquant aux principaux obstacles technologiques aux progrès de la médecine personnalisée. Les cadres de confidentialité et d'éthique permettent une analyse et une évaluation approfondies des systèmes actuels et futurs, garantissant ainsi la pérennité des plateformes utilisées dans le cadre réglementaire actuel et à venir. En établissant des liaisons avec d'autres projets PHRT / SPHN, le projet DPPH cherche à couvrir tous les besoins en Suisse, en commençant par des prototypes à l'échelle nationale. Le projet DPPH garanti également une pertinence et une cohérence internationale, en s'appuyant sur des connexions déjà établies avec la Global Alliance for Genomics and Health (Alliance Globale pour la génomique et la santé, GA4GH) et son groupe de sécurité informatique.