RESEARCH CENTRE

Inria Saclay Center

IN PARTNERSHIP WITH:

Université Versailles Saint-Quentin

2022 ACTIVITY REPORT

Project-Team PETRUS

PErsonal & TRUSted cloud

DOMAIN

Perception, Cognition and Interaction

THEME

Data and Knowledge Representation and Processing

Inria

Contents

Project-Team PETRUS **1**

1 Team members, visitors, external collaborators **2**

2 Overall objectives **2**

3 Research program **3**

4 Application domains **4**
4.1 Personal cloud, home care, IoT, sensing, surveys . . . 4

5 New software and platforms **4**
5.1 New software . . . 4
5.1.1 PlugDB . . . 4
5.2 New platforms . . . 5

6 New results **5**
6.1 PDMS Architecture for Intel SGX (Axis 1) . . . 5
6.2 Privacy-by-design Architecture for Consent-compliant Data Reuse (Axis 1) . . . 6
6.3 DISPERS (extended study)(Axis 2) . . . 6
6.4 Edgelet Computing: Opportunistic Queries on Secure Edges (Axis 3) . . . 7

7 Bilateral contracts and grants with industry **7**
7.1 Bilateral contracts with industry . . . 7
7.2 Bilateral grants with industry . . . 8

8 Partnerships and cooperations **8**
8.1 International research visitors . . . 8
8.1.1 Visits to international teams . . . 8
8.2 National initiatives . . . 8
8.2.1 iPoP, interdisciplinary Project on Privacy, PEPR Cybersécurité (Oct 2022 - Oct 2028) 8
8.2.2 YPPOG, Youth Privacy Protection in Online Gaming, DATAIA project (Sept. 2021 - Sept. 2024) . . . 9
8.2.3 GDP-ERE, GDPR and Personal Cloud: from Empowerment to Accountability, DATAIA project (Sept. 2018 - Jan. 2022) . . . 9

9 Dissemination **10**
9.1 Promoting scientific activities . . . 10
9.1.1 Scientific events: organisation . . . 10
9.1.2 Scientific events: selection . . . 10
9.1.3 Journal . . . 10
9.1.4 Invited talks . . . 10
9.1.5 Scientific expertise . . . 10
9.1.6 Research administration . . . 11
9.2 Teaching - Supervision - Juries . . . 11
9.2.1 Teaching . . . 11
9.2.2 Supervision . . . 11
9.2.3 Juries . . . 12

10 Scientific production **12**
10.1 Major publications . . . 12
10.2 Publications of the year . . . 13

Project-Team PETRUS

Creation of the Project-Team: 2017 July 01

Keywords

Computer sciences and digital sciences

A1.1.8. – Security of architectures

A1.1.9. – Fault tolerant systems

A1.3. – Distributed Systems

A3.1.2. – Data management, quering and storage

A3.1.3. – Distributed data

A3.1.5. – Control access, privacy

A3.1.6. – Query optimization

A3.1.9. – Database

A3.1.11. – Structured data

A4.5. – Formal methods for security

A4.7. – Access control

A4.8. – Privacy-enhancing technologies

Other research topics and application domains

B2.5.3. – Assistance for elderly

B6.4. – Internet of things

B6.6. – Embedded systems

B9.10. – Privacy

1 Team members, visitors, external collaborators

Research Scientists

- Nicolas Anciaux [Team leader, INRIA, Senior Researcher, HDR]
- Luc Bouganim [INRIA, Senior Researcher, HDR]

Faculty Members

- Philippe Pucheral [UVSQ, Professor, HDR]
- Iulian Sandu Popa [UVSQ, Associate Professor, HDR]
- Guillaume Scerri [UVSQ (now ENS Paris Saclay), Associate Professor, until Aug 2022]

PhD Students

- Robin Carpentier [UVSQ]
- Ludovic Javet [INRIA]
- Julien Mirval [Cozy Cloud]

Technical Staff

- Mariem Brahem Habibi [INRIA, Engineer]

Interns and Apprentices

- Pablo Février [Inria, from May 2022 until Oct 2022, master student at INSA-VDL]
- Annicka Ratsirahonana [Inria, from Jun 2022 until Aug 2022, master student at Polytech Sorbonne]
- Ba Tuan Thai [Inria, from Apr 2022 until Sep 2022, master student at PSL]

Administrative Assistant

- Katia Evrat [INRIA]

External Collaborator

- Benjamin Nguyen [INSA CENTRE VDL, HDR]

2 Overall objectives

We are witnessing an exponential accumulation of personal data on central servers: data automatically gathered by administrations and companies but also data produced by individuals themselves (e.g., photos, agendas, data produced by smart appliances and quantified-self devices) and deliberately stored in the cloud for convenience. The net effect is, on the one hand, an unprecedented threat on data privacy due to abusive usage and attacks and, on the other hand, difficulties in providing powerful user-centric services (e.g. personal big data) which require crossing data stored today in isolated silos. The Personal Cloud paradigm holds the promise of a Privacy-by-Design storage and computing platform, where each individual can gather her complete digital environment in one place and share it with applications and users, while preserving her control. However, this paradigm leaves the privacy and security issues in user's hands, which leads to a paradox if we consider the weaknesses of individuals' autonomy in terms of computer security, ability and willingness to administer sharing policies. The challenge is however

paramount in a society where emerging economic models are all based - directly or indirectly - on exploiting personal data.

While many research works tackle the organization of the user's workspace, the semantic unification of personal information, the personal data analytics problems, the objective of the PETRUS project-team is to tackle the privacy and security challenges from an architectural point of view. More precisely, our objective is to help providing a technical solution to the personal cloud paradox. More precisely, our goals are (i) to propose new architectures (encompassing both software and hardware aspects) and administration models (decentralized access and usage control models, data sharing, data collection and retention models) for secure personal cloud data management, (ii) to propose new secure distributed database indexing models, privacy preserving query processing strategies and data anonymization techniques for the personal cloud, and (iii) study economic, legal and societal issues linked to secure personal cloud adoption.

3 Research program

To tackle the challenge introduced above, we identify three main lines of research:

- (Axis 1) Personal cloud server architectures and administration models. Based on the intuition that user control, security and privacy are key properties in the definition of trusted personal cloud solutions, our objective is to propose new architectures (encompassing both software and hardware aspects) for secure personal cloud data management and formally prove important bricks of the architecture. We also focus in this axis on administration models and their enforcement in relation to the architecture of the system, so that the exclusive control of a non expert individual can be ensured.

- (Axis 2) Global query evaluation. The goal of this line of research is to provide capabilities for crossing data belonging to multiple individuals (e.g., performing statistical queries over personal data, computing queries on social graphs or organizing participatory data collection) in a fully decentralized setting while providing strong and personalized privacy guarantees. This means proposing new secure distributed database indexing models and query processing strategies. In addition, we concentrate on locally ensuring to each participant the good behaviour of the processing, such that no collective results can be produced if privacy conditions are not respected by other participants.

- (Axis 3) Technical, legal and economical issues linked to PDMS adoption. This research axis is more transverse and entails multidisciplinary research, addressing the links between economic, legal, societal and technological aspects. We are particularly interested in some specific issues related to the design, implementation and deployment of real PDMS solutions.

Our contributions also rely on tools (algorithms, protocols, proofs, etc.) from other communities, namely security (cryptography, secure multiparty computations, formal methods, differential privacy, etc.) and distributed systems (distributed hash tables, gossip protocols, etc.). Beyond the research actions, we structure our software activity around advanced platforms integrating our main research contributions. These platforms are cornerstones to help validating our research results through accurate performance measurements, a common practice in the DB community, and target the best conferences. It is also a strong vector to federate the team, simplify the bootstrapping of new PhD or master students, conduct multi-disciplinary research and open the way to industrial collaborations and technological transfers. Our main platform is called PlugDB and has reached a high level of maturity. It runs on a microcontroller and is integrated in a real PDMS home box solution deployed in the field for social-medical care. In addition, we are developing a second platform (which is only in the research prototype stage), to provide the user with a PDMS solution that can be hosted on a cloud platform using trusted execution environments (such as Intel SGX) to ensure data privacy and security.

4 Application domains

4.1 Personal cloud, home care, IoT, sensing, surveys

As stated in the software section, the Petrus research strategy aims at materializing its scientific contributions in an advanced hardware/software platform with the expectation to produce a real societal impact. Hence, our software activity is structured around a common Secure Personal Cloud platform rather than several isolated demonstrators. This platform will serve as the foundation to develop a few emblematic applications.

Several privacy-preserving applications can actually be targeted by a Personal Cloud platform, like: (i) smart disclosure applications allowing the individual to recover her personal data from external sources (e.g., bank, online shopping activity, insurance, etc.), integrate them and cross them to perform personal big data tasks (e.g., to improve her budget management) ; (ii) management of personal medical records for care coordination and well-being improvement; (iii) privacy-aware data management for the IoT (e.g., in sensors, quantified-self devices, smart meters); (iv) community-based sensing and community data sharing; (v) privacy-preserving studies (e.g., cohorts, public surveys, privacy-preserving data publishing). Such applications overlap with all the research axes described above but each of them also presents its own specificities. For instance, the smart disclosure applications will focus primarily on sharing models and enforcement, the IoT applications require to look with priority at the embedded data management and sustainability issues, while community-based sensing and privacy-preserving studies demand to study secure and efficient global query processing.

Among these applications domains, one is already receiving a particular attention from our team. Indeed, we gained a strong expertise in the management and protection of healthcare data through our past DMSP (Dossier Medico-Social Partagé) experiment in the field. This expertise is being exploited to develop a dedicated healthcare and well-being personal cloud platform. We are currently deploying 10000 boxes equipped with PlugDB in the context of the DomYcile project. In this context, we are currently setting up an Inria Innovation Lab with the Hippocad company to industrialize this platform and deploy it at large scale (see Section the bilateral contract OwnCare II-Lab).

5 New software and platforms

5.1 New software

5.1.1 PlugDB

Keywords: Databases, Personal information, Privacy, Hardware and Software Platform

Functional Description: PlugDB is a complete platform dedicated to a secure and ubiquitous management of personal data. It aims at providing an alternative to a systematic centralization of personal data. The PlugDB engine is a personal database server capable of storing data (tuples and documents) in tables and BLOBs, indexing them, querying them in SQL, sharing them through assertional access control policies and enforcing transactional properties (atomicity, integrity, durability).

The prototype version of PlugDB engine is embedded in a tamper-resistant hardware device combining the security of smartcard with the storage capacity of NAND Flash. The personal database is hosted encrypted in NAND Flash and the PlugDB engine code runs in the tamper-resistant device. Complementary modules allow to pre-compile SQL queries for the applications, communicate with the DBMS from a remote Java program, synchronize local data with remote servers (typically used for recovering the database in the case of a broken or lost devices) and participate in distributed computation (e.g., global queries). Then, PlugDB was extended to run both on secure devices provided by Gemalto and on specific secure devices designed by PETRUS and assembled by electronic SMEs. Mastering the hardware platform opens up new research and experiment opportunities (e.g., support for wireless communication, secure authentication, sensing capabilities, battery powered ...).

PlugDB engine has been registered first at APP (Agence de Protection des Programmes) in 2009 - a new version being registered every two years - and the hardware datasheets in 2015. PlugDB has been experimented in the field, notably in the healthcare domain. PlugDB was used in an educational platform that we set up : SIPD (Système d'Information Privacy- by-Design). SIPD was used at ENSIIE, INSA CVL and UVSQ through the Versailles Sciences Lab fablab, to raise students awareness of privacy protection problems and embedded programming.

PlugDB combines several research contributions from the team, at the crossroads of flash data management, embedded data processing and secure distributed computations. It then strongly federates all members of our team (permanent members, PhD students and engineers). It is also a vector of visibility, technological transfer and dissemination and gives us the opportunity to collaborate with researchers from other disciplines around a concrete privacy-enhancing platform.

PlugDB is currently industrialized in the context of the OwnCare Inria Innovation Lab (II-Lab). In OwnCare, PlugDB acts as a secure personal cloud to manage medical/social data for people receiving care at home. It is currently being deployed over 10.000 patient in the Yvelines district. The industrialization process covers the development of a complete testing environment, the writing of a detailed documentation and the development of additional features (e.g., embedded ODBC driver, TPM support, flexible access control model and embedded code upgrade notably). It has also required the design of a new hardware platform equipped with a battery power supply, introducing new energy consumption issues for the embedded software.

URL: https://project.inria.fr/plugdb/

Authors: Nicolas Anciaux, Luc Bouganim, Philippe Pucheral, Aydogan Ersoz, Laurent Schneider

Contact: Nicolas Anciaux

5.2 New platforms

Participants: Nicolas Anciaux, Luc Bouganim, Robin Carpentier *(correspondent)*, Iulian Sandu Popa.

Personal Data Management Systems (PDMS) arrive at a rapid pace boosted by smart disclosure initiatives and new regulations such as GDPR. However, our recent survey [1] indicates that the existing PDMS solutions cover partially the PDMS data life-cycle and, more importantly, focus on specific privacy threats depending on the employed architecture. To address this issue, we proposed in [1] a logical reference architecture for an extensive (i.e., covering all the major functionalities) and secure (i.e., circumventing all the threats specific to the PDMS context) PDMS. We also discussed several possible physical instances fo the architecture and showed that TEEs (Trusted Execution Environments) are a prime option for building a trustworthy PDMS platform [2].

Hence, based on our previous studies, we have developed a first prototype of an extensive and secure PDMS (ES-PDMS) platform using the state-of-the-art TEE technology available today, i.e., Intel Software Guard eXtension (SGX). The originality of our approach is to achieve extensibility through a set of isolated data-oriented tasks potentially untrusted by the PDMS owner, running alongside a trusted module which controls the complete workflow and limits data leakage [14]. Our ES-PDMS software stack can be deployed on any SGX-enabled machine (i.e., any relatively recent computer having an Intel CPU). This prototype was presented in a demonstration paper [15] focusing on security properties of the platform with the help of several concrete scenarios and interactive games.

6 New results

6.1 PDMS Architecture for Intel SGX (Axis 1)

Participants: Nicolas Anciaux, Luc Bouganim, Robin Carpentier, Iulian Sandu Popa *(correspondent)*.

Personal Data Management Systems (PDMSs) arrive at a rapid pace providing individuals with appropriate tools to collect, manage and share their personal data. At the same time, the emergence of Trusted Execution Environments (TEEs) opens new perspectives in solving the critical and conflicting challenge of securing users' data while enabling a rich ecosystem of data-driven applications. Hence, our approach is to propose a PDMS architecture leveraging TEEs as a basis for security (see section 5.2). Unlike existing solutions, our architecture allows for data processing extensiveness through the integration of any user-defined functions, albeit untrusted by the data owner. In this context, we focused on aggregate computations of large sets of database objects and provided a first study to mitigate the very large potential data leakage [14]. We introduce the necessary security building blocks and show that an upper bound on data leakage can be guaranteed to the PDMS user. We then proposed in [13] practical evaluation strategies ensuring that the potential data leakage remains minimal with a reasonable performance overhead. Finally, we validated our proposal with an Intel SGX-based PDMS implementation on real data sets and presented a demonstration of the platform in [15]. This work is part of Robin Carpentier's PhD [22, subcite].

6.2 Privacy-by-design Architecture for Consent-compliant Data Reuse (Axis 1)

Participants: Nicolas Anciaux, Mariem Brahem *(correspondent)*, Guillaume Scerri.

Participatory sensing allows individuals to contribute data across time and space in order to feed general interest computation tasks. However, there are major obstacles including the preservation of the privacy of contributors. This consideration has led to two main approaches, sometimes combined, which are, respectively, to trade privacy for rewards, and to take advantage of privacy-enhancing technologies anonymizing the collected data. Although relevant, we claim that these approaches do not sufficiently take into account the consent of the participants to the use of the personal data provided and may even lead to *defects in consent* even in the presence of a trusted system. To address this issue, we introduce the ℓ-completeness property, which ensures that the data provided can be reused for all the tasks to which their contributors consent as long as they are analyzed with a same set of $\ell - 1$ other sources, without causing any defect in consent. We propose a privacy by design implementation of the solution leveraging trusted execution environment (Intel SGX), to enforce this property. This study, which was conducted in collaboration with Valerie Issarny (Inria Mimove team), appears in 2022 at the Pervasive and Mobile Computing (PMC) journal [12], and extends our previous study in [4].

6.3 DISPERS (extended study) (Axis 2)

Participants: Luc Bouganim *(correspondent)*, Iulian Sandu Popa.

Personal Data Management System (PDMS) solutions are flourishing, boosted by smart disclosure initiatives and new regulations. PDMSs allow users to easily store and manage data directly generated by their devices or resulting from their (digital) interactions. Users can then leverage the power of their PDMS to benefit from their personal data, for their own good and in the interest of the community. The PDMS paradigm thus brings exciting perspectives by unlocking novel usages, but also raises security issues. An effective approach, considered in several recent works, is to let the user data distributed on personal platforms, secured locally using hardware and/or software security mechanisms. This work goes beyond the local security issues and addresses the important question of securely querying this massively distributed personal data. In this work, we propose DISPERS (DIStributed Privacy-presERving querieS) which can: (i) be applied to very large P2P systems; (ii) select specific data source nodes to achieve query

pertinence; (iii) reach a lower-bound on leakage in the presence of colluding nodes with *near-certainty*, i.e., with a very high and adjustable probability; (iv) adjust the trade-off between security level and security cost. Thus, DISPERS enables generic, efficient and scalable P2P data computations while still providing users with strong confidentiality guarantees. We design DISPERS in an incremental way by considering increasingly stronger threat models. For each of them, we derive a *security requirement* that must be satisfied to reach the lower-bound on leakage. More precisely, we consider: (a) *spied nodes* when the attacker can spy on PDMS communications, leading to the **H**idden communication requirement and the $DISPERS^{H}$ protocol in which communications are protected through encryption and anonymization; (b) *leaking nodes* when the attacker can, additionally, observe the internal states of the PDMS, leading to the **R**andom dispersion of data requirement and the $DISPERS^{HR}$ protocol in which data-at-rest, i.e. the indexes, are protected using Shamir secret sharing and data-in-use are protected through task compartmentalization; and finally, (c) *corrupted nodes* when the attacker fully controls some PDMSs (thus can additionally alter its behavior), leading to the **C**ollaborative proofs requirement and to the $DISPERS^{HRC}$ protocol, in which a contributing or indexing node answers a request for sensitive data only if it obtains a collaborative proof that the request is justified. This work was published in the VLDB Journal [11] and is based on previous studies: in [9], we showed that the execution of a P2P query can indeed rely exclusively on data processor nodes if and only if they are selected in a verifiable random way, which cannot be influenced by corrupted nodes.

6.4 Edgelet Computing: Opportunistic Queries on Secure Edges (Axis 3)

Participants: Nicolas Anciaux *(correspondent)*, Luc Bouganim, Ludovic Javet, Philippe Pucheral.

We call Edgelet computing the current convergence between Opportunistic Network (OppNet) and Trusted Execution Environment (TEE) at the very edge of the network. We believe that this convergence bears the seeds of a novel and important class of applications leveraging fully decentralized and highly secure computations among data scattered on multiple personal devices. In this work, we tackle the data management and distributed system issues related to this environment : we characterize the Edgelet computing paradigm by a combination of architectural and computing assumptions, an unusual threat model and three properties that guarantee the liveness, safety and security of executions; we propose two alternative execution strategies enforcing liveness in opposite ways and discuss their impact on safety; we provide preliminary quantitative and qualitative evaluations of these strategies and derives from them design rules to adapt centralized computations to the Edgelet computing paradigm. This work has been published in the CCGRID'22 conference [17]. Then, a prototype has been developed and a demonstration paper, accepted in PERCOM'23 [16], highlights the pertinence of the approach through a real medical use-case. We currently try to extend the applicability of the solution beyond the OppNet context, by considering TEE-enabled devices communicating through various forms of degraded channels.

7 Bilateral contracts and grants with industry

7.1 Bilateral contracts with industry

OwnCare-2 IILab (Jan 2022 - Dec 2025) - Partners: PETRUS, Hippocad

Participants: Philippe Pucheral *(correspondent)*, Laurent Schneider, Nicolas Anciaux, Luc Bouganim.

The OwnCare IILab – Inria Innovation Lab - (Jan 2018-Dec 2021) aimed at conceiving a secured personal medical folder facilitating the organization of medical and social care provided at home to elderly people and at deploying it in the field. This IILab has been built in partnership with the Hippocad company

which won, in association with Inria and UVSQ, a public call for tender launched by the Yvelines district to deploy this medical folder on the whole distinct (10.000 patients). This solution, named DomYcile in the Yvelines district, is based on a home box combining the PlugDB hardware/software technology developed by the Petrus team (to manage and secure the medical folder) and additional technology developed by Hippocad. The primary result of the OwnCare IILab has been to build a concrete industrial solution based on PlugDB and deploy it so far among 3000 patients in the Yvelines district, despite the Covid pandemia. In 2022, Hippocad has become a subsidiary of the La Poste group opening new opportunities in terms of deployment. Hence, Inria, UVSQ and Hippocad have launched a follow up of the OwnCare IILab for the period Jan 2022-Dec 2025. The goal of the OwnCare2 IILab is (1) to integrate our solution in the MaSanté 2022 national roadmap by making it interoperable with external services (without hurting the security provided by the box), (2) to handle, in a privacy-preserving way, new usages like actimetrics, teleassistance and global statistics based on IoT techniques, machine learning and decentralized computations and (3) try to deploy it at the national/international level.

7.2 Bilateral grants with industry

Cozy Cloud CIFRE - Mirval contract (Nov 2020 - Oct 2023) - Partners: Cozy Cloud, PETRUS

Participants: Julien Mirval *(correspondent)*, Iulian Sandu Popa, Luc Bouganim.

A third CIFRE PhD thesis has been started between Cozy Cloud and Julien Mirval from PETRUS. Cozy Cloud is a French startup providing a personal Cloud platform. The Cozy product is a software stack that anyone can deploy to run his personal server in order to host his personal data and web services. The objective of this thesis is to propose appropriate solutions to effectively train an AI model (e.g., a deep neural network) in a fully distributed system while providing strong security guarantees to the participating nodes. The results, in the form of protocols and distributed and secure execution algorithms, will be applied to practical cases provided by the Cozy Cloud company.

8 Partnerships and cooperations

8.1 International research visitors

8.1.1 Visits to international teams

Research stays abroad

- Iulian Sandu Popa did a one month secondment (April 29 to May 28, 2022) visit to the Municipality of Thira, Santorini as part of the H2020 MASTER (Multiple ASpect TrajEctoRy management and analysis) project. During his stay, he had scientific meetings with Nikos Pelekis (UPRC) and Zaineb Chelly Dagdia (UVSQ) to start research collaborations on topics related to massive multi-aspect trajectory data management and their privacy, and applying data mining solutions to multi-aspect trajectory data and related security issues respectively.

8.2 National initiatives

8.2.1 iPoP, interdisciplinary Project on Privacy, PEPR Cybersécurité (Oct 2022 - Oct 2028)

Partners: Inria, CNRS, EDHEC, INSA CVL, INSA Lyon, UGA, Université de Lille, Université Rennes 1, UVSQ, CNIL

Digital technologies provide services that can greatly increase quality of life (e.g. connected e-health devices, location based services or personal assistants). However, these services can also raise major privacy risks, as they involve personal data, or even sensitive data. Indeed, this notion of personal data

is the cornerstone of French and European regulations, since processing such data triggers a series of obligations that the data controller must abide by. This raises many multidisciplinary issues, as the challenges are not only technological, but also societal, judiciary, economic, political and ethical. The objectives of this project are thus to study the threats on privacy that have been introduced by these new services, and to conceive theoretical and technical privacy-preserving solutions that are compatible with French and European regulations, that preserve the quality of experience of the users. These solutions will be deployed and assessed, both on the technological and legal sides, and on their societal acceptability. In order to achieve these objectives, we adopt an interdisciplinary approach, bringing together many diverse fields: computer science, technology, engineering, social sciences, economy and law.

The project's scientific program focuses on new forms of personal information collection, on the learning of Artificial Intelligence (AI) models that preserve the confidentiality of personal information used, on data anonymization techniques, on securing personal data management systems, on differential privacy, on personal data legal protection and compliance, and all the associated societal and ethical considerations. This unifying interdisciplinary research program brings together internationally recognized research teams (from universities, engineering schools and institutions) working on privacy, and the French Data Protection Authority (CNIL).

8.2.2 YPPOG, Youth Privacy Protection in Online Gaming, DATAIA project (Sept. 2021 - Sept. 2024)

Partners: CERDI (Université Paris Saclay), LITEM (IMT-BS), PETRUS (Inria-UVSQ).
Despite its somewhat seemingly light nature, Youth Privacy Protection in Online Gaming is an very important topic in the field of Privacy. Indeed, 94% of minor children (under 18 years old) play video games, and 60

As stated above, our approach is multi-disciplinary, mixing computer science, to propose privacy preserving algorithms and infrastructures to manage youth online data, economy in order to understand the constraints of the gaming industry and how to take them into account when it comes to legal aspects, and finally law in order to propose procedures to help the companies move to legally compliant processes.

One example of a technical prototype, developed by PETRUS in 2022 (see the git repository) is a wrapper for the publicly available data of one of the most played games online, League of Legends (LoL). We are currently setting up a field study to see if it is possible to train AI algorithms to detect, based on this open data, if a player is under or over 18.

8.2.3 GDP-ERE, GDPR and Personal Cloud: from Empowerment to Accountability, DATAIA project (Sept. 2018 - Jan. 2022)

Partners: DANTE (U. of Versailles), PETRUS (Inria-UVSQ).
The role of individuals and the control of their data is a central issue in the new European regulation (GDPR) enforced on 25th May 2018. Data portability is a new right provided under those regulations. It allows citizens to retrieve their personal data from the companies and governmental agencies that collected them, in an interoperable digital format. The goals are to enable the individual to get out of a captive ecosystem, and to favor the development of innovative personal data services beyond the existing monopolistic positions. The consequence of this new right is the design and deployment of technical platforms, commonly known as Personal Cloud. But personal cloud architectures are very diverse, ranging from cloud based solutions where millions of personal cloud are managed centrally, to self-hosting solutions. These diversity is not neutral both in terms of security and from the point of view of the chain of liabilities. The GDP-ERE project tends to study those issues in an interdisciplinary approach by the involvement of jurists and computers scientists. The two main objectives are (i) to analyze the effects of the personal cloud architectures on legal liabilities, enlightened by the analysis of the rules provided under the GDPR and (ii) to propose legal and technological evolutions to highlight the share of liability between each relevant party and create adapted tools to endorse those liabilities.

9 Dissemination

9.1 Promoting scientific activities

9.1.1 Scientific events: organisation

Member of the organizing committees

- Luc Bouganim: Co-organizer "École thématique BDA Masses de Données Distribuées", Bastia, 2022

9.1.2 Scientific events: selection

Member of the conference program committees

- Nicolas Anciaux: BDA'22.
- Luc Bouganim: BDA'22.
- Iulian Sandu Popa: SSDBM'22, DATA'22, IEEE Mobile Cloud'22.

9.1.3 Journal

Member of the editorial boards

- Nicolas Anciaux: Associate Editor at the VLDB Journal (until sept 2022).

Reviewing activities

- Nicolas Anciaux: Pervasive and Mobile Computing.

9.1.4 Invited talks

- Nicolas Anciaux: Séminaire SoSySec, "Security and privacy in personal data management systems", organized by IRISA, Jan. 2022.
- Iulian Sandu Popa: invited talk at Journées GDR Sécurité Infomatique, June 24, 2022.

9.1.5 Scientific expertise

- Nicolas Anciaux: vice-president of the Inria CRCN-ISFP recruitment committee 2022 for Inria Saclay research center.
- Nicolas Anciaux: member of the jury of the 7th edition of CNIL-Inria Privacy Award 2022
- Luc Bouganim: Member of the selection committee (COS) for the position of enseignants-chercheurs contractuels" - INSA CVL.
- Luc Bouganim: Member of the selection committee (COS) for the position of "professeur" - UVSQ.
- Luc Bouganim: CIFRE evaluation for ANRT
- Philippe Pucheral: Member of the selection committee (COS) for the position of "professeur" - Univ. Paris-Nanterre.
- Philippe Pucheral: Member of the selection committee (COS) for the position of "Associate Professor" - Univ. Toulouse 3.

9.1.6 Research administration

- Nicolas Anciaux: deputy scientific director (DSA) of Inria Saclay research center.
- Nicolas Anciaux: member of the CoDiReV of Univ. Paris-Saclay (committee of the research heads of the Univ. Paris-Saclay components and partners).
- Nicolas Anciaux: member of the Research Commission of the ACademic Council (CR-CAC) of Univ. Paris-Saclay.
- Nicolas Anciaux: Member of the Bureau of David lab at UVSQ.
- Luc Bouganim: Member of the Scientific Commission (CS) of Inria Saclay-IDF (Cordi-S, Post-Doc, Delegation).
- Luc Bouganim: PhD thesis referent for the Doctoral School of University Paris-Saclay
- Philippe Pucheral: Elected member in the council of the 'Computer Sciences' Graduate School of Paris-Saclay University.
- Iulian Sandu Popa: Member of the Technological Development Commission (CDT) at Inria de Saclay (ADT funding and SED engineer requests).
- Iulian Sandu Popa: Member of the Bureau of David lab at UVSQ.

9.2 Teaching - Supervision - Juries

9.2.1 Teaching

- Philippe Pucheral: head of the M1 and M2 DataScale master program at University Paris-Saclay.
- Master: Iulian Sandu Popa, Bases de données relationnelles (niveau M1), Gestion des données spatiotemporelles (niveau M2), Sécurité des bases de données (niveau M2), 96, UVSQ, France. Philippe Pucheral, courses in M1 and M2 in databases and in security, introductory courses for jurists,UVSQ, France.
- Licence: Iulian Sandu Popa, Bases de données (niveau L2 et L3), 96, UVSQ, France.
- Engineers school: Nicolas Anciaux, Databases (ENSTA, module IN206, M1), 32, and Advanced databases (ENSTA, module ASI13, level M2), 32. Luc Bouganim, Bases de données relationnelles (ENSTA, module IN207, M1), 32.
- MOOC : "Villes intelligentes : défis technologiques et sociétaux". Co-Auteurs: Nicolas Anciaux, Stéphane Grumbach, Valérie Issarny, Nathalie Mitton, Christine Morin, Animesh Pathak et Hervé Rivano. Sessions sur la plateforme FUN en 2022.

9.2.2 Supervision

- Defended PhD (Decembre 14, 2022): Robin Carpentier, Privacy-preserving third-party computations on secure personal data management systems [22, subcite], UVSQ, Nicolas Anciaux and Iulian Sandu Popa.
- PhD in progress: Julien Mirval, DISSEC-ML : DIStributed and SECure Machine Learning on Personal Clouds, UVSQ, November 2020, Luc Bouganim and Iulian Sandu Popa.
- PhD in progress: Ludovic Javet, Requêtes distribuées respectueuses de la vie privée dans un environnement partiellement connecté, Inria, Januray 2020, Luc Bouganim, Nicolas Anciaux and Philippe Pucheral.

9.2.3 Juries

- Nicolas Anciaux: Reviewer and member of the PhD Jury of Jean-Yves Zie Diali (INSA CVL Bourges, 21/06/2022).

10 Scientific production

10.1 Major publications

[1] N. Anciaux, P. Bonnet, L. Bouganim, B. Nguyen, P. Pucheral, I. Sandu-Popa and G. Scerri. 'Personal Data Management Systems: The security and functionality standpoint'. In: *Information Systems* 80 (2019), pp. 13–35. DOI: 10.1016/j.is.2018.09.002. URL: https://hal.archives-ouvertes.fr/hal-01898705.

[2] N. Anciaux, L. Bouganim, P. Pucheral, I. S. Popa and G. Scerri. 'Personal Database Security and Trusted Execution Environments: A Tutorial at the Crossroads'. In: *Proceedings of the VLDB Endowment (PVLDB)* (Aug. 2019). DOI: 10.14778/3352063.3352118. URL: https://hal.inria.fr/hal-02269292.

[3] L. Bouganim, J. Loudet and I. Sandu Popa. 'Highly distributed and privacy-preserving queries on personal data management systems'. In: *The VLDB Journal*. Online first articles (7th July 2022). DOI: 10.1007/s00778-022-00753-1. URL: https://hal.inria.fr/hal-03814840.

[4] M. Brahem, G. Scerri, N. Anciaux and V. Issarny. 'Consent-driven data use in crowdsensing platforms: When data reuse meets privacy-preservation'. In: PerCom 2021 - IEEE International Conference on Pervasive Computing and Communications. Kassel / Virtual, Germany, 22nd Mar. 2021. URL: https://hal.inria.fr/hal-03097047.

[5] H. Comon, C. Jacomme and G. Scerri. 'Oracle simulation: a technique for protocol composition with long term shared secrets'. In: ACM CCS 2020. CCS '20: Proceedings of the 2020 ACM SIGSAC Conference on Computer and Communications Security. Orlando, United States: Association for Computing Machinery, 9th Nov. 2020, pp. 1427–1444. URL: https://hal.inria.fr/hal-02913866.

[6] L. Javet, N. Anciaux, L. Bouganim and P. Pucheral. 'Edgelet Computing: Pushing Query Processing and Liability at the Extreme Edge of the Network'. In: CCGrid 2022. Taormina, Italy, 16th May 2022. URL: https://hal.inria.fr/hal-03666895.

[7] R. Ladjel, N. Anciaux, P. Pucheral and G. Scerri. 'Trustworthy Distributed Computations on Personal Data Using Trusted Execution Environments'. In: *TrustCom 2019 - The 18th IEEE International Conference on Trust, Security and Privacy in Computing and Communications / BigDataSE 2019 - 13th IEEE International Conference on Big Data Science and Engineering*. Rotorua, New Zealand, Aug. 2019. DOI: 10.1109/TrustCom/BigDataSE.2019.00058. URL: https://hal.archives-ouvertes.fr/hal-02269207.

[8] S. Lallali, N. Anciaux, I. Sandu-Popa and P. Pucheral. 'Supporting secure keyword search in the personal cloud'. In: *Information Systems* 72 (Dec. 2017), pp. 1–26. DOI: 10.1016/j.is.2017.09.003. URL: https://hal.inria.fr/hal-01660599.

[9] J. Loudet, I. Sandu-Popa and L. Bouganim. 'SEP2P: Secure and Efficient P2P Personal Data Processing'. In: *EDBT 2019 - 22nd International Conference on Extending Database Technology*. Lisbon, Portugal, Mar. 2019. URL: https://hal.inria.fr/hal-01949641.

[10] S. J. Pan, I. Sandu-Popa and C. Borcea. 'DIVERT: A Distributed Vehicular Traffic Re-Routing System for Congestion Avoidance'. In: *IEEE Transactions on Mobile Computing* 16.1 (Jan. 2017), pp. 58–72. DOI: 10.1109/TMC.2016.2538226. URL: https://hal.inria.fr/hal-01426424.

10.2 Publications of the year

International journals

[11] L. Bouganim, J. Loudet and I. Sandu Popa. 'Highly distributed and privacy-preserving queries on personal data management systems'. In: *The VLDB Journal*. Online first articles (7th July 2022). DOI: 10.1007/s00778-022-00753-1. URL: https://hal.inria.fr/hal-03814840.

[12] M. Brahem, G. Scerri, N. Anciaux and V. Issarny. 'Consent-driven Data Reuse in Multi-tasking Crowdsensing Systems: A Privacy-by-Design Solution'. In: *Pervasive and Mobile Computing* 83 (2022). DOI: 10.1016/j.pmcj.2022.101614. URL: https://hal.archives-ouvertes.fr/hal-03775759.

International peer-reviewed conferences

[13] R. Carpentier, I. Sandu Popa and N. Anciaux. 'Data Leakage Mitigation of User-Defined Functions on Secure Personal Data Management Systems'. In: SSDBM 2022 - 34th International Conference on Scientific and Statistical Database Management. Copenhagen, Denmark, 6th July 2022. DOI: 10.1145/3538712.3538741. URL: https://hal.inria.fr/hal-03692175.

[14] R. Carpentier, I. Sandu Popa and N. Anciaux. 'Local Personal Data Processing with Third Party Code and Bounded Leakage'. In: DATA 2022 - 11th International Conference on Data Science, Technology and Applications. Lisbon, Portugal, 11th July 2022. URL: https://hal.inria.fr/hal-03686098.

[15] R. Carpentier, F. Thiant, I. Sandu Popa, N. Anciaux and L. Bouganim. 'An Extensive and Secure Personal Data Management System Using SGX'. In: EDBT 2022 - 25th International Conference on Extending Database Technology. Edinburgh / Virtual, United Kingdom, 29th Mar. 2022. URL: https://hal.inria.fr/hal-03580286.

[16] L. Javet, N. Anciaux, L. Bouganim, L. Lamoureux and P. Pucheral. 'Secure Computations in Opportunistic Networks: An Edgelet Demonstration with a Medical Use-Case'. In: PerCom 2023 - 21st IEEE International Conference on Pervasive Computing and Communications. Atlanta, United States, 13th Mar. 2023. URL: https://hal.inria.fr/hal-03925679.

[17] L. Javet, N. Anciaux, L. Bouganim and P. Pucheral. 'Edgelet Computing: Pushing Query Processing and Liability at the Extreme Edge of the Network'. In: CCGrid 2022. Taormina, Italy, 16th May 2022. URL: https://hal.inria.fr/hal-03666895.

Conferences without proceedings

[18] R. Carpentier, I. S. Popa and N. Anciaux. 'Data Leakage Mitigation of User-Defined Functions on Secure Personal Data Management Systems'. In: BDA 2022 - 38ème Conférence sur la Gestion de Données - Principes, Technologie et Applications. Clermont-Ferrand, France, 24th Oct. 2022. URL: https://hal.inria.fr/hal-03941879.

[19] R. Carpentier, F. Thiant, I. S. Popa, N. Anciaux and L. Bouganim. 'An Extensive and Secure Personal Data Management System Using SGX'. In: BDA 2022 - 38ème Conférence sur la Gestion de Données - Principes, Technologie et Applications. Clermont-Ferrand, France, 24th Oct. 2022. DOI: 10.48786/edbt.2022.53. URL: https://hal.inria.fr/hal-03941868.

[20] L. Javet, N. Anciaux, L. Bouganim and P. Pucheral. 'Edgelet Computing: Pushing Query Processing and Liability at the Extreme Edge of the Network'. In: BDA 2022 - 38ème Conférence sur la Gestion de Données – Principes, Technologies et Applications. Clermont-Ferrand, France, 24th Oct. 2022. URL: https://hal.inria.fr/hal-03925714.

[21] L. Javet, N. Anciaux, L. Bouganim and P. Pucheral. 'Edgelet Computing: Pushing Query Processing and Liability at the Extreme Edge of the Network'. In: APVP 2022 - 12ème Atelier sur la Protection de la Vie Privée. Châtenay-sur-Seine, France, 13th June 2022. URL: https://hal.inria.fr/hal-03925653.

Doctoral dissertations and habilitation theses

[22] R. Carpentier. 'Privacy-preserving third-party computations on secure personal data management systems'. Université Paris-Saclay, 14th Dec. 2022. URL: https://hal.inria.fr/tel-03912320.

Other scientific publications

[23] R. Carpentier, F. Thiant, I. Sandu Popa, N. Anciaux and L. Bouganim. *Outline: An Extensive and Secure Personal Data Management System Using SGX*. Paris, France, 28th June 2022. URL: https://hal.inria.fr/hal-03763815.