RESEARCH CENTRE

Saclay - Île-de-France

IN PARTNERSHIP WITH:

Université Versailles Saint-Quentin

2021
ACTIVITY REPORT

Project-Team
PETRUS

PErsonal & TRUSted cloud

DOMAIN

Perception, Cognition and Interaction

THEME

Data and Knowledge Representation and Processing

Contents

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 [Univ de Versailles Saint-Quentin-en-Yvelines, Professor, HDR]
- Iulian Sandu Popa [Univ de Versailles Saint-Quentin-en-Yvelines, Associate Professor, HDR]
- Guillaume Scerri [Univ de Versailles Saint-Quentin-en-Yvelines, Associate Professor]

Post-Doctoral Fellows

- Mariem Habibi [Inria]
- Riad Ladjel [Inria, until Jun 2021]

PhD Students

- Robin Carpentier [Univ de Versailles Saint-Quentin-en-Yvelines]
- Ludovic Javet [Inria]
- Julien Mirval [Cozy Cloud, CIFRE]

Technical Staff

- Laurent Scheider [inria, Engineer]
- Floris Thiant [Inria, Engineer]

Interns and Apprentices

- Abdelhafid Belhabib [Inria, from May 2021 until Sep 2021]
- Billal Zemmoura [Univ de Versailles Saint-Quentin-en-Yvelines, from May 2021 until Sep 2021]

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) Economic, legal and societal issues. This research axis is more transverse and entails multidisciplinary research, addressing the links between economic, legal, societal and technological aspects. We will follow here a multi-disciplinary approach based on a 3-step methodology: i) identifying important common issues related to privacy and to the exploitation of personal data; ii) characterizing their dimensions in all relevant disciplines and jointly study their entanglement; iii) validating the proposed analysis, models and trade-offs thanks to in vivo experiments.

These contributions will 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 a single common platform (rather than isolated demonstrators), integrating our main research contributions, called PlugDB. This platform is the cornerstone to help validating our research results through accurate performance measurements on a real platform, 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.

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: Floris Thiant *(correspondent)*, Nicolas Anciaux, Luc Bouganim, Iulian Sandu Popa, Guillaume Scerri.

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 currently develop 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). Our ES-PDMS software stack can be deployed on any SGX-enabled machine (i.e., any relatively recent computer having an Intel CPU). We have acquired a server with 6 Intel Xeon E-2276G CPU cores allowing a collaborative development of the ES-PDMS protoype by the Petrus team as well as the related Personal Cloud applications. We note that the Intel Xeon CPU series offer access to the Intel Attestation Service[1] which is required for the remote enclave attestation process. This allows us to implement

[1] See Intel website

(and not only to simulate) remote attestations required by several use-cases in the PDMS context (e.g., distributed computations between a network of PDMSs or attesting results of local computations on the PDMS to a third party).

6 New results

6.1 PDMS Architecture for Intel SGX (Axis 1)

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

To design a platform for ES-PDMS [1], our approach is to use trusted execution environments. The challenge is to orchestrate the different data-related tasks using secure hardware enclaves, in order to provide rich data-oriented functionality while meeting the desired security objectives. We have begun the implementation of a new PDMS platform based on Intel SGX (see section 5.2). By decomposing the computation into data tasks, we aim to guarantee the *bilateral security* property, on the one hand by protecting the data confidentiality and privacy of the PDMS owner, and on the other hand by guaranteeing the integrity of the computed results shared with third parties. This work is part of Robin Carpentier's PhD. A short paper describing the underlying research problem was presented at EuroS&P 2021 [14]. A demonstration illustrating the proposal has also been designed.

6.2 Consent-driven Data Reuse (Axis 1)

Participants: Mariem Brahem *(correspondent)*, Nicolas Anciaux, 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. We propose a clustering strategy sensitive to the data distribution, which is shown to optimize data reuse and utility. This study [13], conducted in collaboration with Valerie Issarny (Inria Mimove team), appears at PerCom 2021. The design of an participatory sensing architecture leveraging on trusted execution environments to support stronger attackers models is ongoing. A journal version of the paper is currently submitted to a special issue of PMC journal following PerCom 2021, proposing a SGX implementation of the solution, and therefore allowing for a fully malicious service provider.

6.3 Secure Mobile Participatory Sensing (Axis 2)

Participants: Iulian Sandu Popa *(correspondent)*.

Mobile participatory sensing (MPS) could benefit many application domains. A major domain is smart transportation, with applications such as vehicular traffic monitoring, vehicle routing, or driving behavior analysis. However, MPS's success depends on finding a solution for querying large numbers of smart phones or vehicular systems, which protects user location privacy and works in real-time. This work proposes PAMPAS, a privacy-aware mobile distributed system for efficient data aggregation in MPS. In PAMPAS, mobile devices enhanced with secure hardware, called secure probes (SPs), perform distributed query processing, while preventing users from accessing other users' data. A supporting server infrastructure (SSI) coordinates the inter-SP communication and the computation tasks executed on SPs. PAMPAS ensures that SSI cannot link the location reported by SPs to the user identities even if SSI has additional background information. Moreover, an enhanced version of the protocol, named PAMPAS+, makes the system robust even against advanced hardware attacks on the SPs. Hence, the risk of user location privacy leakage remains very low even for an attacker controlling the SSI and a few corrupted SPs. Our experimental results demonstrate that these protocols work efficiently on resource constrained SPs being able to collect the data, aggregate them, and share statistics or derive models in real-time. This work [12] has been accomplished in collaboration with NJIT and DePaul University.

6.4 DISSEC-ML (Axis 2)

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

A Personal Data Management System (PDMS) allows its owner to easily collect, store and manage data, directly generated by her devices, or resulting from her interactions with companies or administrations. PDMSs unlock innovative usages by crossing multiple data sources from one or many users, thus requiring aggregation primitives. Indeed, aggregation primitives are essential to compute statistics on user data and are also a fundamental building block for machine learning algorithms. Typically, CozyCloud (the PETRUS partner for Julien's CIFRE) wishes to use distributed aggregation for a Naive Bayes classifier. In this work, we study a protocol allowing for secure aggregation in a massively distributed PDMS environment, which adapts to selective participation and PDMSs characteristics, and is reliable with respect to failures, with no compromise on accuracy. Initial results were published in [15] and [16] with preliminary experiments showing the effectiveness of our protocol. It can adapt to several contexts with varying PDMSs characteristics in terms of communication speed or CPU resources and can adjust the aggregation strategy to the estimated selective participation.

6.5 Edgelet Computing: Opportunistic Queries on Secure Edges (Axis 2)

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 was submitted recently and we are now focusing more specifically on machine learning algorithms within an Edgelet computing environment.

6.6 Hidding Communications Patterns in Distributed Queries (Axis 2)

Participants: Riad Ladjel *(correspondent)*, Nicolas Anciaux, Guillaume Scerri.

The execution of distributed queries on populations of PDMS involves communication patterns between computing nodes (see [6]), which may depend on the values of the personal data being processed (a computing node may aggregate personal data corresponding to a given range of sensitive values). An attacker observing communications, even encrypted, can potentially infer personal information about participants. The use of traditional solutions to conceal data-dependent communications at runtime results in either significant performance penalties or privacy gains that are difficult to quantify. Chapter 4 of Riad Ladel's PhD manuscript formulates as an ϵ-differential privacy problem the issue of concealing communication patterns in distributed queries. In collaboration with Aurélien Bellet (Inria Magnet team), we proposed algorithms using the differential confidentiality model (ϵ,δ)-differential privacy, allowing a finer analysis of the trade-offs between privacy, performance and utility in practice (see [18]).

6.7 Empowerment and Big Data on Personal Data (Axis 3)

Participants: Nicolas Anciaux *(correspondent)*, Riad Ladjel.

The place of individuals and the control of their data have emerged as central issues in the European data protection regulation. The *empowerment* of the individual has notably resulted in the recognition of a new prerogative for the individual: the right to the portability of personal data. The corollary of this new right is the design and deployment of technical platforms, commonly known as Personal Data Mangement Servers (PDMS) or PIMS, allowing the individual to consolidate all his or her data in a single system managed under his or her control. On the strength of these technical and legal innovations, several questions arise: what forms of empowerment are targeted in practice? What are the appropriate conditions to guarantee the objective pursued? At the crossroads of these questions, one dimension appears to be insufficiently exploited: that of *agency*. During this period, in collaboration with law researchers Célia Zolynski (IRJS, Sorbonne Univ.) and Sébastien Chaudat (DANTE, Univ. of Versailles), we have transposed this notion from the social sciences to the management of personal data and provided a new reading of the empowerment measures of Big Data functionalities on personal data. This study led in 2021 to a journal publication in the Global Privacy Law Review [11].

7 Bilateral contracts and grants with industry

7.1 Bilateral contracts with industry

OwnCare II-Lab (Jul 2017 - Dec 2021) Partners: PETRUS, Hippocad
End 2016, the Yvelines district lauched a public call for tender to deploy a personal medical-social folder aiming at covering the whole distinct (10.000 patients). The Hippocad company, in partnership with Inria, won this call for tender with a solution called DomYcile and the project was launched in July 2017. DomYcile is based on a home box combining the PlugDB hardware/software technology developed by the Petus team and a communication layer based on SigFox. Hippocad and Petrus then decided to launch a joint II-Lab (Inria Innovation Lab) named OwnCare in January 2018. The objective is threefold: (1) build an industrial solution based on PlugDB and deploy it in the Yvelines district in the short-term, (2) use this Yvelines testbed to improve the solution and try to deploy it at the national/international level in the medium-term and (3) design flexible/secure/mobile personal medical folder solutions addressing new usages (IoT, machine learning models, distributed statistics, etc.) in the long-term. In 2021, important efforts have been put (1) on the deployment of the boxes in the field despite the Covid pandemia, (2) on pursuing the design of decentralized privacy-preserving distributed computations and (3) on initiating research work related to daily activity discovery thanks to home sensors and machine learning algorithms. On the administrative part, a followup of the II-Lab for the next 4 years is under preparation.

7.2 Bilateral grants with industry

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

8.1.1 ANR PerSoCloud (Jan 2017 - Mar 2021)

Partners: Orange Labs (coordinator), PETRUS (Inria-UVSQ), Cozy Cloud, U. of Versailles.
The objective of PerSoCloud is to design, implement and validate a full-fledged Privacy-by-Design Personal Cloud Sharing Platform. One of the major difficulties linked to the concept of personal cloud lies in organizing and enforcing the security of the data sharing while the data is no longer under the control of a central server. We identify three dimensions to this problem. Devices-sharing: assuming that the primary copy of user U1's personal data is hosted in a secure place, how to share and synchronize it with U1's multiple (mobile) devices without compromising security? Peers-sharing: how user U1 could exchange a subset of his-her data with an identified user U2 while providing to U1 tangible guarantees about the usage made by U2 of this data? Community-sharing: how user U1 could exchange a subset of his-her data with a large community of users and contribute to personal big data analytics while providing to U1 tangible guarantees about the preservation of his-her anonymity? In addition to tackling these three scientific and technical issues, a legal analysis will guarantee compliance of this platform with the security and privacy French and UE regulation, which firmly promotes the Privacy by Design principle, including the current reforms of personal data regulation.

8.1.2 GDP-ERE, 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.

8.2 Regional initiatives

8.2.1 Prevadom, Paris-Saclay grant (2020-2021)

Partners: Inria (PETRUS).
The objective of the Prevadom project is to integrate IoT facilities and machine learning algorithms into the DomYcile PlugDB-enabled boxes to detect and ideally prevent loneliness and despair situations. Such situations are peculiarly alarming during periods where patients are confined at home (e.g., during a pandemy). To this end, the PlugDB-enabled boxes will be augmented with IoT communication facilities in order to collect data from sensors (e.g., light, temperature, hygrometry, motion) and quantified-self

devices (e.g., thermometer, oxymeter, blood pressure). Such raw data are highly intrusive. The objective is then to store them and analyse them with machine learning models locally (i.e., inside the box) and only externalize aggregated data or launch alarms when required.

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 (cancelled in 2021, trying in 2022)

Member of the conference program committees

- Nicolas Anciaux: Int. Conf. on Extending Database Technology (EDBT 2021), IEEE International Conference on Smart Computing (SmartComp 2021)
- Luc Bouganim: Int. Conf. on Extending Database Technology (EDBT 2021).
- Iulian Sandu Popa: ACM Int. Conf. on Data Management (Sigmod 2022), Int. Conf. on Scientific and Statistical Database Management (SSDBM 2021), Int. Conf. on Data Science, Technology and Applications (DATA 2021), IEEE Mobile Cloud 2021, Workshop on Fairness, Accountability, Transparency, Ethics and Society on the Web (FATES 2021 - joint with The Web Conference 2021)

Member of the editorial boards

- Nicolas Anciaux: Associate Editor at the VLDB Journal
- Luc Bouganim: Member of editorial board for the special issue (BDA 2020 best papers) in TLDKS Journal.

Reviewer - reviewing activities

- Nicolas Anciaux: PMC Journal
- Iulian Sandu Popa: ACM SIGMOD Record, International Journal of Geo-Information

9.1.2 Invited talks

- Nicolas Anciaux: Seminar, "Personal Database Management Systems (PDMS) : vers une plateforme de Big Data citoyen ?", Groupe de Travail Protection de la Vie Privée (GT-PVP), May 2021

9.1.3 Scientific expertise

- Nicolas Anciaux: member of the jury of the 6th edition of CNIL-Inria Privacy Award.

9.1.4 Research administration

- Nicolas Anciaux: Vice-head of sciences (DSA) at Inria Saclay (since Jul 2021)
- Nicolas Anciaux: Member of CAC, CR CAC, CodireV at Université Paris-Saclay (since end of 2021)
- Nicolas Anciaux: Inria Saclay representative at "Formation par la Recherche (FpR)" and "Mission Jeunes Chercehurs (MJC)" (until Jul 2021)
- Nicolas Anciaux: Member of the "COnseil de la POlitique Doctorale (COPOD)" at University Paris-Saclay (until Jul 2021)

- Nicolas Anciaux: correspondent for the STIC doctoral school at Inria-Saclay (until Jul 2021)
- Nicolas Anciaux: Member of the "Bureau du Comité des Projets (BCEP)" at Inria-Saclay
- Nicolas Anciaux: Member of the "Bureau du Laboratoire" at DAVID Lab
- 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: Member of the council of the « Computer Science » Graduate School of University Paris-Saclay
- Philippe Pucheral: Member of the 'Habilitation à Diriger des Recherches' committee of University of Versailles - UFR des Sciences
- Iulian Sandu Popa: Member of *Commission du Développement Technologique* (CDT) at Inria-Saclay
- Iulian Sandu Popa: Member of the *Bureau du Laboratoire* at DAVID Lab at UVSQ
- Iulian Sandu Popa: PhD thesis referent for the Doctoral School of University Paris-Saclay

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. Guillaume Scerri, Sécurité et bases de données pour juristes, courses in databases and security (M1 and M2), 48 hours, UVSQ, France.
- Licence: Guillaume Scerri, Databases, systems, 126 hours, UVSQ, France. Iulian Sandu Popa, Bases de données (niveau L2 et L3), 96, UVSQ, France.
- Engineers school : Nicolas Anciaux, Databases (module IN206, niveau M1), 30, and Advanced databases (module ASI13, niveau M2), 30. Luc Bouganim, Bases de données relationnelles (niveau M1), 32, ENSTA, France, Philippe Pucheral, Gestion de Données Avancées (niveau M1), 30, ENSIIE Evry, France.
- MOOC : "Villes intelligentes : défis technologiques et sociétaux", organisé par Valérie Issarny et Nathalie Mitton. Co-Auteurs: Nicolas Anciaux, Stéphane Grumbach, Valérie Issarny, Nathalie Mitton, Christine Morin, Animesh Pathak et Hervé Rivano. Sessions sur la plateforme FUN.

9.2.2 Supervision

- PhD in progress: Robin Carpentier, Secure and efficient data processing in trusted execution environments for the personal cloud, UVSQ, October 2018, Nicolas Anciaux, Iulian Sandu Popa and Guillaume Scerri
- 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 and selection committees

- Nicolas Anciaux : Member of the jury for the recruitment of researchers at Inria Saclay (CRCN-IFSP 2021, Inria Saclay)
- Nicolas Anciaux : Reviewer of the HDR of Mathieu Cunche (INSA Lyon, 2/6/2021)
- Nicolas Anciaux : Reviewer of the HDR of Vasile Radu Ciucanu (INSA Centre Val de Loire, 9/6/2021)
- Nicolas Anciaux : Member of the PhD jury of Paul Alexandre Marillonnet (Institut Polytechnique de Paris, 30/11/2021)
- Luc Bouganim: Tutor of the HDR of Iulian Sandu Popa (UVSQ, 10/12/2021)
- Luc Bouganim: Member of the selection committee (COS) for the position of "enseignants-chercheurs contractuels" - INSA CVL.
- Luc Bouganim: Member of the recruitment committee of associate professor at ITU Copenhagen.
- Iulian Sandu Popa: Member of the selection committee for ATER (*Attaché temporaire d'enseignement et de recherche*) at UVSQ (June 2021).

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] N. Anciaux and C. Zolynski. 'Empowerment et Big Data sur données personnelles : de la portabilité à l'agentivité'. In: *Le Big Data et le Droit*. Thèmes et Commentaire. Dalloz, 2020. URL: https://hal.inria.fr/hal-02349274.

[4] C. Berthet, C. Zolynski, N. Anciaux and P. Pucheral. '" Contenus numériques et récupération des données : un nouvel outil d' 'empouvoirement' du consommateur ? "' In: *Dalloz IP/IT* IP IT / 10 (Jan. 2017). URL: https://hal.inria.fr/hal-01429939.

[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] 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.

[7] 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.

[8] 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.

[9] 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] P. Tran-Van, N. Anciaux and P. Pucheral. ‘SWYSWYK: A Privacy-by-Design Paradigm for Personal Information Management Systems’. In: *International Conference on Information Systems Development (ISD)*. Cyprus, Cyprus, Sept. 2017. URL: https://hal.inria.fr/hal-01675090.

10.2 Publications of the year

International journals

[11] N. Anciaux, C. Zolynski, S. Chaudat and R. Ladjel. ‘Empowerment and Big Personal Data: from Portability to Personal Agency’. In: *Global Privacy Law Review* (2021). URL: https://hal.inria.fr/hal-03140409.

[12] I. Sandu Popa, D. H. T. That, K. Zeitouni and C. Borcea. ‘Mobile participatory sensing with strong privacy guarantees using secure probes’. In: *Geoinformatica* 25.3 (July 2021), pp. 533–580. DOI: 10.1007/s10707-019-00389-4. URL: https://hal.archives-ouvertes.fr/hal-03329908.

International peer-reviewed conferences

[13] 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.

[14] R. Carpentier, I. S. Popa and N. Anciaux. ‘Poster: Reducing Data Leakage on Personal Data Management Systems’. In: EuroS&P 2021 - 6th IEEE European Symposium on Security and Privacy. Vienne, Austria, 6th Sept. 2021. DOI: 10.1109/EuroSP51992.2021.00057. URL: https://hal.inria.fr/hal-03536381.

[15] J. Mirval, L. Bouganim and I. Sandu Popa. ‘Practical Fully-Decentralized Secure Aggregation for Personal Data Management Systems’. In: 33rd International Conference on Scientific and Statistical Database Management, SSDBM 2021. SSDBM 2021: 33rd International Conference on Scientific and Statistical Database Management, Tampa, FL, USA, July 6-7, 2021. Tampla, FL, United States, 2021, pp. 259–264. DOI: 10.1145/3468791.3468821. URL: https://hal.archives-ouvertes.fr/hal-03329878.

Conferences without proceedings

[16] J. Mirval, L. Bouganim and I. Sandu Popa. ‘Practical Fully-Decentralized Secure Aggregation for Personal Data Management Systems’. In: BDA 2021 - 37ème Conférence sur la Gestion de Données - Principes, Technologies et Applications. Paris, France, 25th Oct. 2021. URL: https://hal.archives-ouvertes.fr/hal-03538834.

Doctoral dissertations and habilitation theses

[17] I. Sandu Popa. ‘Preserving Individual Privacy with Personal Data Management Systems’. Université de Versailles Saint-Quentin-en-Yvelines; Université Paris-Saclay, 10th Dec. 2021. URL: https://hal.inria.fr/tel-03531619.

Reports & preprints

[18] R. Ladjel, N. Anciaux, A. Bellet and G. Scerri. *Mitigating Leakage from Data Dependent Communications in Decentralized Computing using Differential Privacy*. 24th Dec. 2021. URL: https://hal.inria.fr/hal-03502320.