RESEARCH CENTRES

Inria Centre at Université Grenoble Alpes

Inria Lyon Centre

IN PARTNERSHIP WITH:

Institut national des sciences appliquées de Lyon

2024 ACTIVITY REPORT

Project-Team PRIVATICS

Privacy Models, Architectures and Tools for the Information Society

IN COLLABORATION WITH: Centre d'innovation en télécommunications et intégration de services

DOMAIN

Algorithmics, Programming, Software and Architecture

THEME

Security and Confidentiality

Inria

Contents

Project-Team PRIVATICS

Creation of the Project-Team: 2014 July 01

Keywords

Computer sciences and digital sciences

A1.2.5. – Internet of things

A4.8. – Privacy-enhancing technologies

A5.1.9. – User and perceptual studies

A9.9. – Distributed AI, Multi-agent

Other research topics and application domains

B6.3.1. – Web

B6.3.2. – Network protocols

B9.6.2. – Juridical science

B9.10. – Privacy

1 Team members, visitors, external collaborators

Research Scientists

- Vincent Roca [Team leader, INRIA, Researcher]
- Nataliia Bielova [INRIA, Senior Researcher, from Oct 2024]
- Nataliia Bielova [INRIA, Researcher, until Sep 2024]
- Claude Castelluccia [INRIA, Senior Researcher]
- Heber Hwang Arcolezi [INRIA, ISFP]
- Cedric Lauradoux [INRIA, Researcher]
- Mohamed Maouche [INRIA, ISFP]
- Cristiana Teixeira Santos [INRIA, Starting Research Position, from Oct 2024 until Oct 2024]
- Cristiana Teixeira Santos [INRIA, Starting Research Position, from Jun 2024 until Jul 2024]
- Cristiana Teixeira Santos [INRIA, Starting Research Position, from Apr 2024 until Apr 2024]

Faculty Members

- Antoine Boutet [INSA LYON, Associate Professor]
- Mathieu Cunche [INSA LYON, Professor]
- Clementine Gritti [INSA LYON, from Jun 2024]

Post-Doctoral Fellows

- Sanju Ahuja [INRIA, Post-Doctoral Fellow, from Nov 2024]
- Karel Kubicek [INRIA, Post-Doctoral Fellow, from Dec 2024]
- Alexandre Lodie [INRIA, Post-Doctoral Fellow, until Jun 2024]
- Abhishek Mishra [INRIA, Post-Doctoral Fellow, from Oct 2024]
- Abhishek Mishra [INSA LYON, Post-Doctoral Fellow, until Sep 2024]

PhD Students

- Jan Aalmoes [INSA LYON, until Nov 2024]
- Ivan Baheux-Blin [MURENA SAS, CIFRE, from Feb 2024]
- Teodora Curelariu [UGA]
- Thomas Lebrun [INRIA, until Nov 2024]
- Jules Marmier [INRIA, from Nov 2024]
- Gilles Mertens [INRIA]
- Alix Ntoutoume Nzame [OPEN SEZAM SAS, CIFRE]
- Samuel Pelissier [INSA LYON, until Sep 2024]

Technical Staff

- Mohamed Bechorfa [INRIA, Engineer]
- Nathan Brunet [INRIA, Engineer, until May 2024]
- Hugo Dabadie [INRIA, Engineer, until Jun 2024]
- Lucas Magnana [INRIA, Engineer, from Jun 2024]

Interns and Apprentices

- Marouane Akassab [INRIA, Intern, from May 2024 until Jul 2024]
- Nicolas Audermatte [INRIA, Intern, from May 2024 until Jul 2024]
- Elodie Bernard [INSA LYON, Intern, from Feb 2024 until Jul 2024]
- Matthew Bunch [INRIA, Intern, from May 2024 until Jul 2024]
- Mohamed Amine Dahmouni [CNRS, Intern, from Apr 2024 until Sep 2024]
- Zakaria El Kazdam [INRIA, Intern, from Jun 2024 until Aug 2024]
- Paul Retourne [INRIA, Intern, from Jul 2024 until Jul 2024]
- Arno Venaille [INRIA, Intern, from Jun 2024 until Aug 2024]
- Helain Zimmermann [INRIA, Intern, from Jun 2024 until Aug 2024]

Administrative Assistant

- Helen Pouchot-Rouge-Blanc [INRIA]

Visiting Scientists

- Kumari Kancherla [IIT Gandhinagar, from Apr 2024 until Jun 2024]
- Juliet Samandari [UNIV BROWN, from Oct 2024 until Nov 2024]

2 Overall objectives

2.1 Context and overall objectives

From ambient privacy to massive and ubiquitous data collection: In a very short span of time, we switched from a world where "ambient privacy" was the rule, to a situation dominated by massive, ubiquitous and precise data collections, where trying to protect our privacy requires constant efforts. If, 50 years ago, the perceived threat was that of an *state surveillance* (e.g., the SAFARI project led to the creation in 1978 of the French privacy regulation and the DPA, CNIL), nowadays, *"capitalism surveillance"*, a term popularized by Shoshana Zubboff, is a concern of equal, if not greater, importance. It has been made possible by the super-fast development of the Web in the 1990s, of smartphones ten years later, and now of IoT devices of all kinds, and all these technological breakthroughs led to the creation of highly profitable giant companies, most of which leverage on user-data for profiling and targeting.

Undoubtedly, this digital world opened major opportunities, highly beneficial to the society in general and to individuals in particular. However, it also poses considerable privacy threats that can potentially turn these new technologies into a nightmare if they are not accompanied by appropriate legal and ethical rules. As the French "Loi Informatique et Liberté" (1978) says in its first chapter: "Information technology must be at the service of every citizen. [...] It must not infringe on human identity, human rights, privacy, or personal or public freedom."

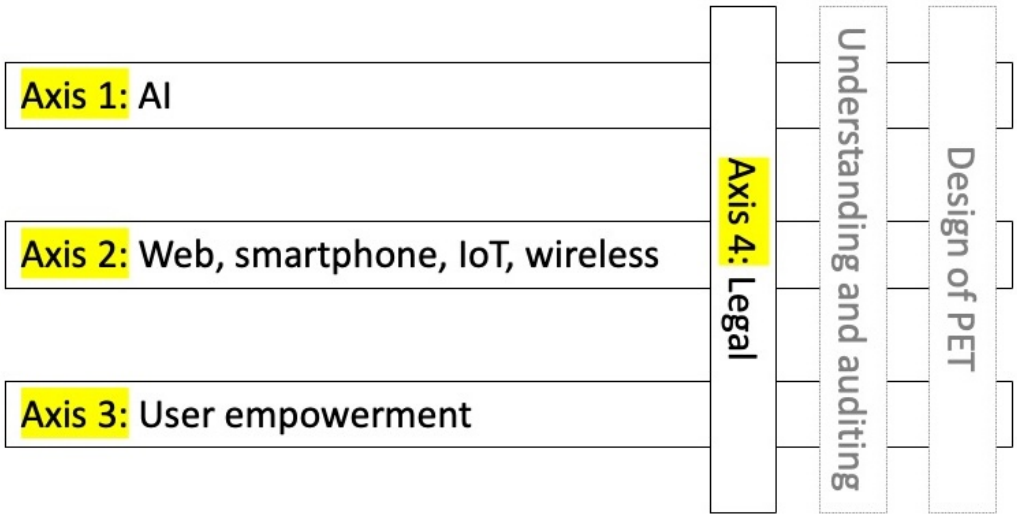
Making the world – a little bit – better: Privacy is thus essential to protect individuals, for instance against potential misuses of personal data. Privacy is also essential to protect the society, as has been highlighted by the misuse of personal data in order to surreptitiously influence voters in elections (e.g., Cambridge Analytica). But privacy is too important to be left only in the hands of individuals: the role of regulators and Data Protection Authorities is fundamental from this viewpoint, leading to regulations (e.g., GDPR) that protect all citizens by default.

In this landscape, public research has a key role to play. By working in a complementary manner, from highly theoretical subjects up to the reverse engineering of deployed systems, or the design of privacy enhancing technologies, public research also contributes to making the world – a little bit – better.

3 Research program

PRIVATICS activities: Since its creation in 2014, the PRIVATICS team focuses on privacy protection in this digital world, and its members contribute to the domain through theoretical, practical, but also transdisciplinary activities. Indeed, while the team mainly focuses on technical aspects of privacy, the team also interacts with legal, economical dimension of privacy. In order to be impactful, for our research community but also for the society, the approach followed is fundamentally transdisciplinary. It covers the computer-science, legal and design domains, with sometimes sociological contributions, by the means of enriched collaborations with the members of these disciplines.

4 Application domains



More specifically, our activities cover four main research axes, depicted above, namely:

1. the **"AI" research** axis includes works on "privacy considerations in ML" (e.g., Federated ML and the explainability of Automated Decision Systems), but also on the "use of ML for privacy" (e.g., for medical report anonymisation);

2. the **"Web, smartphone, IoT and wireless networks"** (e.g., BLE and LoRaWAN) research axis focuses on several types of connected devices and services, responsible of major data leaks, for which our contributions can be highly impactful. We conducted large scale measurements, we reverse-engineered several technologies, and we proposed Privacy Enhancement Technologies (PET) when appropriate;

3. the **"User Empowerment"** research axis studies how users keep control over their data and how they are being manipulated. For example, this axis involves large-scale measurement of consent

on the Web (in form of cookie banners), dark patterns that manipulate users' decision making when interacting with consent, and tensions with legal requirements for GDPR consent when designing consent banners – this axes is particularly advanced, at the intersection with the "Legal" axis presented below.

4. the **"Legal"** research axis intersects all previous axes, and consists in transdisciplinary research in Computer Science and Law. We analyze *legal requirements for compliance with the EU Data Protection Laws* of systems and services, such as cookie banners, providers of such banners and their legal roles and responsibilities (e.g., we refined legal high-level requirements into concrete system requirements, such as 22 low-level requirements to assess compliance of consent banners). We also analyze the technical and regulation aspects of privacy invasive technologies that present significant risks (e.g., face recognition, or intelligent surveillance cameras). In front of such complex problems having both technical and legal dimensions, advances are only possible through a transdisciplinary work with legal scholars.

Across these topics, we work on:

- the analysis of systems and services in order to *understand them, sometimes to audit them* (e.g., by measuring personal data leaks through large scale measurement campaigns);
- the design of privacy enhancement technologies (PET), in various domains (e.g., to reduce privacy risks in wireless technologies, or to enhance privacy properties of Federated ML).

Transdisciplinarity made concrete: Privacy being fundamentally at the crossroad of several domains, many hard research questions that we address in the previous four research axes, require a transdisciplinary approach, where experts of different domains share their expertise and benefit from one another. This is the approach we deliberately chose. Therefore, PRIVATICS works with scholars in the legal, economist, design, and social science domains. It takes various forms: participation to common funded projects (e.g., the IPoP and CovOMM projects), participation to common research activities, co-direction of legal PhDs, recruitment of a legal Post-Doctorate, recruitment of an Inria International Chair Junior, and publications in legal venues. PRIVATICS makes transdisciplinarity concrete.

5 Social and environmental responsibility

5.1 Environmental impacts of research results

The activities of PRIVATICS are not directly related to environmental considerations. However, promoting privacy in a connected world advocates for less data collection and processing, as opposed to massive data collection and big data. From this point of view, we believe that our research results are aligned with environmental considerations.

5.2 Societal impacts of research results

Collaborating with regulators thanks to an independent expertise: Developing an *independent expertise* is part of our values. Although big tech companies, such as GAFA, contribute to several privacy enhancing technologies (e.g., in the AI domain) and can offer funding opportunities, we chose not to go into that direction.

We believe that the most efficient way to combat the "surveillance capitalism" doctrine these companies created is to work with regulators. France since 1978 with the "Loi Informatique et Liberté", the EU with the privacy regulation (GDPR, ePrivacy, DMA/DSA) during the past years, and now with AI regulation, paved the way for a better world, more respectful of the individual human rights, internationally. We contribute concretely to this trend.

Since the beginning, PRIVATICS works closely with the French Data Protection Agency, CNIL. During the period, it took the form of a temporary leave to CNIL for Nataliia Bielova, the nomination of Claude Castelluccia as a CNIL commissioner, the participation of CNIL in the IOTics and now IPoP projects, and feedback to several CNIL and EDPB public consultations. Our work is also cited in legal decisions (Belgian

DPA). Additionally, several PRIVATICS members are experts for ENISA (Claude Castelluccia and Cédric Lauradoux), member of ENISA Data Protection Engineering Working Group (Claude Castelluccia), EDPB (Mathieu Cunche, Nataliia Bielova) and EU Commission for the implementaiton of the DSA (Nataliia Bielova). We contributed to several landmark reports on these topics.

We believe that PRIVATICS successfully helped regulators during this period, bringing our expertise at various levels in various ways. We think this is the best approach to be impactful, in particular with respect to giant Internet companies whose business model is so profitable that they have little incentives to change it, unless obliged to do so.

Contributing to the establishment of doctrines regarding technologies that can have major societal impacts: Certain new technologies raise major questions, with societal and ethical potential implications. We contributed to the establishment of doctrines through reports on AI regulation, facial recognition regulation. The major involvement of Claude Castelluccia in the CNIL Board enabled to contribute to the French DPA doctrine with respect to various important subjects related to new technologies. Being in position to influence public doctrines, independently of any private interest, following a scientific approach, is part of the major outcomes of our team.

Transfer to well chosen private companies and public administrations: Our decision not to work with GAFA does not imply we do not work with private companies. We have two future CIFRE PhD with small French companies in the domain of online, sovereign identity management, and "de-googleized" operating system for smartphones. We have several projects with public hospitals and administrations. We provided technical expertise (under confidentiality clauses) on data protection for Europol, the French and German ministry of the interior about the implementation of the EPRIS framework in the proposal for a regulation on automated data exchange for police cooperation ("Prüm II"). Those collaborations open the way for concrete and mutually beneficial transfers, in line with our values.

Contributing to international standards: Being able to contribute to standards in order to promote our views and research outputs, is a highly efficient manner to have concrete impacts. One of us did it for privacy extensions of IEEE 802 standards and was officially recognized for his expertise. In a totally different domain, another one co-chaired an IETF group and published 15 RFCs over the time.

Actions towards the general public: Scientific outreach towards the general public is one of our missions, and we significantly contributed. The MOOC on privacy in the digital world attracted a bit more than 40,000 persons, has been qualified "of public interest" by one of the participants. It is one example. Additionally, every year we contribute to the "Fête de la Science" by proposing mini-conferences and working sessions with the young public, one of us regularly goes into high school to promote science and privacy, we participate also to conferences with non-scientific public and we are interviewed by journalists. We take it to heart to "vulgarize" and help our fellow citizens understand this highly complex domain, with so many societal implications. Although we sometimes would like to do more, this is time consuming and we try to find a good balance.

Actions in support of public authorities: In addition to working with regulators (see above), helping public authorities is also part of our missions. We did it during the COVID19 crisis, and our decisive

work on contact and presence tracing protocols, in the context of the public/private StopCovid project-team, contributed to a successful "crisis application". We also contributed, with all the member of this StopCovid project-team, to strengthen the technological and digital sovereignty of the Nation, with a solution focused on the health authority, respectful of our values and choices.

Participation in ethical committees: Additionally, several PRIVATICS members are part of various ethical committees:

- Vincent Roca is member of the Inria COERLE (commité d'évaluation des risques légaux et éthiques);
- Cédric Lauradoux represents the Inria COERLE (commité d'évaluation des risques légaux et éthiques) in the Grenoble research center, helping local researchers to fill in their application form;
- Cédric Lauradoux is member of the University of Grenoble Alps (UGA) ethical committee;
- Mathieu Cunche is member of *Comité d'éthique de la recherche (CER)* of Lyon University.

6 Highlights of the year

6.1 PhD and HDR defenses

Three PhD Students defended their doctoral thesis this year. In chronological order:

- **Samuel Pelissier**, "Privacy-preserving communications for the IoT", INSA de Lyon, Septempber 27th, 2024. Thesis co-supervised by M. Cunche and V. Roca. [32]
- **Thomas Lebrun**, "Health Data: Exploring Emerging Privacy Enhancing Mechanisms", INSA de Lyon, December 5th, 2024. Thesis co-supervised by A. Boutet and M. Cunche. [31]
- **Jan Aalmoes**, "IA pour des services moraux : concilier équité et confidentialité", INSA de Lyon, December 13th, 2024. Thesis co-supervised by A. Boutet and M. Cunche.

Additionally, an HDR defense took place.

- **Antoine Boutet**, "Privacy issues in AI and geolocation: from data protection to user awareness", Habilitation à Diriger des Recherches (HDR), INSA de Lyon, December 10th, 2024. [30]

7 New software, platforms, open data

7.1 New software

7.1.1 gtm-eye

Name: gtm-eye: a browser extension to inspect Google Tag Manager (GTM) configurations

Keyword: Security and Privacy in Web Services

Functional Description: Gtm-eye is a browser extension that enables a user to inspect Google Tag Manager (GTM) configurations, identifying client-side (to some extent server-side) tags, enabling a selective execution of these tags, and displaying scripts loaded by these tags and scripts in a recursive manner.

Release Contributions: First registered version of gtm-eye.

Contact: Vincent Roca

7.1.2 NLP Privacy

Name: NLP Privacy

Keywords: Privacy, Natural language processing

Scientific Description: Associated publication: G. BERTHELIER, A. BOUTET, A. RICHARD, "Toward training NLP models to take into account privacy leakages", in : BigData 2023 - IEEE International Conference on Big Data, IEEE, p. 1–9, Sorrento, Italy, December 2023, [hal:hal-04299405].

Functional Description: This library provides tools to evaluate three privacy risks on NLP models trained on sensitive data: 1) the counterfactual memorization, which corresponds to rare and sensitive information which has too much influence on the model, 2) the membership inference, and 3) the ability to extract verbatim training data from models.

URL: https://gitlab.inria.fr/aboutet1/NLP-Privacy

Contact: Antoine Boutet

7.1.3 nlp-mem

Name: nlp-mem

Keywords: Privacy, NLP, LLM

Functional Description: This lib aims to quantify the memorization of sensitive information by nlp models.

URL: https://gitlab.inria.fr/hdabadie/nlp-attacks

Contact: Antoine Boutet

7.1.4 ldp-audit

Name: Local Differential Privacy Auditor

Keyword: Differential privacy

Functional Description: A tool for auditing Locally Differentially Private (LDP) protocols.

URL: https://github.com/hharcolezi/ldp-audit

Contact: Heber Hwang Arcolezi

8 New results

8.1 Research axis 1: AI

8.1.1 Revealing the True Cost of Locally Differentially Private Protocols: An Auditing Perspective

Participants: Héber Arcolezi, et al..

While the existing literature on Differential Privacy (DP) auditing predominantly focuses on the centralized model (e.g., in auditing the DP-SGD algorithm), we advocate for extending this approach to audit Local DP (LDP). To achieve this, we introduce the LDP-Auditor framework for empirically estimating the privacy loss of locally differentially private mechanisms. This approach leverages recent advances in designing privacy attacks against LDP frequency estimation protocols. More precisely, through the analysis of numerous state-of-the-art LDP protocols, we extensively explore the factors influencing

the privacy audit, such as the impact of different encoding and perturbation functions. Additionally, we investigate the influence of the domain size and the theoretical privacy loss parameters (epsilon) and (delta) on local privacy estimation. In-depth case studies are also conducted to explore specific aspects of LDP auditing, including distinguishability attacks on LDP protocols for longitudinal studies and multidimensional data. Finally, we present a notable achievement of our LDP-Auditor framework, which is the discovery of a bug in a state-of-the-art LDP Python package. Overall, our LDP-Auditor framework as well as our study offer valuable insights into the sources of randomness and information loss in LDP protocols. These contributions collectively provide a realistic understanding of the local privacy loss, which can help practitioners in selecting the LDP mechanism and privacy parameters that best align with their specific requirements. We open-sourced LDP-Auditor in .

Related publication: [7]

8.1.2 On the Impact of Multi-dimensional Local Differential Privacy on Fairness

Participants: Héber Arcolezi, et al..

Automated decision systems are increasingly used to make consequential decisions in people's lives. Due to the sensitivity of the manipulated data and the resulting decisions, several ethical concerns need to be addressed for the appropriate use of such technologies, particularly fairness and privacy. Unlike previous work, which focused on centralized differential privacy (DP) or on local DP (LDP) for a single sensitive attribute, in this paper, we examine the impact of LDP in the presence of several sensitive attributes (i.e., multi-dimensional data) on fairness. Detailed empirical analysis on synthetic and benchmark datasets revealed very relevant observations. In particular, (1) multi-dimensional LDP is an efficient approach to reduce disparity, (2) the variant of the multi-dimensional approach of LDP (we employ two variants) matters only at low privacy guarantees (high), and (3) the true decision distribution has an important effect on which group is more sensitive to the obfuscation. Last, we summarize our findings in the form of recommendations to guide practitioners in adopting effective privacy-preserving practices while maintaining fairness and utility in machine learning applications.

Related publication: [8]

8.1.3 On the Alignment of Group Fairness with Attribute Privacy

Participants: Jan Aalmoes, Antoine Boutet, et al..

Machine learning (ML) models have been adopted for applications with high-stakes decision-making like healthcare and criminal justice. To ensure trustworthy ML models, the new AI regulations (e.g., AI Act) have established several pillars such as privacy, safety and fairness that model design must take into account. Designing such models requires an understanding of the interactions between fairness definitions with different notions of privacy. Specifically, the interaction of group fairness (i.e., protection against discriminatory behaviour across demographic subgroups) with attribute privacy (i.e., resistance to attribute inference attacks-AIAs), has not been comprehensively studied. In this paper, we study in depth, both theoretically and empirically, the alignment of group fairness with attribute privacy in a blackbox setting. We first propose AdaptAIA, which outperforms existing AIAs on real-world datasets with class imbalances in sensitive attributes. We then show that group fairness theoretically bounds the success of AdaptAIA, which depends on the choice of fairness metrics (e.g., demographic parity or equalized odds). Through our empirical study, we show that attribute privacy can be achieved from group fairness at no additional cost other than the already existing trade-off with utility. Our work has several implications: i) group fairness acts as a defense against AIAs, which is currently lacking, ii) practitioners do not need to explicitly train models for both fairness and privacy to meet regulatory requirements, iii) Adap-tAIA can be used for blackbox auditing of group fairness.

Related publication: [11]

8.1.4 Synthetic Data: Generate Avatar Data on Demand

Participants: Thomas Lebrun, Antoine Boutet, Mohamed Maouche, et al..

Anonymization is crucial for the sharing of personal data in a privacy-aware manner yet it is a complex task that requires to set up a trade-off between the robustness of anonymization (i.e., the privacy level provided) and the quality of the analysis that can be expected from anonymized data (i.e., the resulting utility). Synthetic data has emerged as a promising solution to overcome the limits of classical anonymization methods while achieving similar statistical properties to the original data. Avatar-based approaches are a specific type of synthetic data generation that rely on local stochastic simulation modeling to generate an avatar for each original record. While these approaches have been used in healthcare, their attack surface is not well documented and understood. In this paper, we provide an extensive assessment of such approaches and comparing them against other data synthesis methods. We also propose an improvement based on conditional sampling in the latent space, which allows synthetic data to be generated on demand (i.e., of arbitrary size). Our empirical analysis shows that avatar-generated data are subject to the same utility and privacy trade-off as other data synthesis methods with a privacy risk more important on the edge data, which correspond to records that have the fewest alter egos in the original data.

Related publication: [16]

8.1.5 Nob-MIAs: Non-biased Membership Inference Attacks Assessment on Large Language Models with Ex-Post Dataset Construction

Participants: Héber Arcolezi, et al..

The rise of Large Language Models (LLMs) has triggered legal and ethical concerns, especially regarding the unauthorized use of copyrighted materials in their training datasets. This has led to lawsuits against tech companies accused of using protected content without permission. Membership Inference Attacks (MIAs) aim to detect whether specific documents were used in a given LLM pretraining, but their effectiveness is undermined by biases such as time-shifts and n-gram overlaps. This paper addresses the evaluation of MIAs on LLMs with partially inferable training sets, under the ex-post hypothesis, which acknowledges inherent distributional biases between members and non-members datasets. We propose and validate algorithms to create “non-biased” and “non-classifiable” datasets for fairer MIA assessment. Experiments using the Gutenberg dataset on OpenLamma and Pythia show that neutralizing known biases alone is insufficient. Our methods produce non-biased ex-post datasets with AUC-ROC scores comparable to those previously obtained on genuinely random datasets, validating our approach. Globally, MIAs yield results close to random, with only one being effective on both random and our datasets, but its performance decreases when bias is removed.

Related publication: [14]

8.1.6 Let Them Drop: Scalable and Efficient Federated Learning Solutions Agnostic to Stragglers

Participants: Clémentine Gritti, et al..

Secure Aggregation (SA) stands as a crucial component in modern Federated Learning (FL) systems, facilitating collaborative training of a global machine learning model while protecting the privacy of individual clients' local datasets. Many existing SA protocols described in the FL literature operate synchronously, leading to notable runtime slowdowns due to the presence of stragglers (i.e. latearriving clients). To address this challenge, one common approach is to consider stragglers as client failures and use SA solutions that are robust against dropouts. While this approach indeed seems to work, it

unfortunately affects the performance of the protocol as its cost strongly depends on the dropout ratio and this ratio has increased significantly when taking stragglers into account. Another approach explored in the literature to address stragglers is to introduce asynchronicity into the FL system. Very few SA solutions exist in this setting and currently suffer from high overhead. In this paper, similar to related work, we propose to handle stragglers as client failures but design SA solutions that do not depend on the dropout ratio so that an unavoidable increase on this metric does not affect the performance of the solution. We first introduce Eagle, a synchronous SA scheme designed not to depend on the client failures but on the online users' inputs only. This approach offers better computation and communication costs compared to existing solutions under realistic settings where the number of stragglers is high. We then propose Owl, the first SA solution that is suitable for the asynchronous setting and once again considers online clients' contributions only. We implement both solutions and show that: (i) in a synchronous FL with realistic dropout rates (taking potential stragglers into account), Eagle outperforms the best SA solution, namely Flamingo, by ×4; (ii) In the asynchronous setting, Owl exhibits the best performance compared to the state-of-the-art solution LightSecAgg .

Related publication: [26]

8.1.7 A Systematic and Formal Study of the Impact of Local Differential Privacy on Fairness: Preliminary Results

Participants: Héber Arcolezi, et al..

Machine learning (ML) algorithms rely primarily on the availability of training data, and, depending on the domain, these data may include sensitive information about the data providers, thus leading to significant privacy issues. Differential privacy (DP) is the predominant solution for privacy-preserving ML, and the local model of DP is the preferred choice when the server or the data collector are not trusted. Recent experimental studies have shown that local DP can impact ML prediction for different subgroups of individuals, thus affecting fair decision-making. However, the results are conflicting in the sense that some studies show a positive impact of privacy on fairness while others show a negative one. In this work, we conduct a systematic and formal study of the effect of local DP on fairness. Specifically, we perform a quantitative study of how the fairness of the decisions made by the ML model changes under local DP for different levels of privacy and data distributions. In particular, we provide bounds in terms of the joint distributions and the privacy level, delimiting the extent to which local DP can impact the fairness of the model. We characterize the cases in which privacy reduces discrimination and those with the opposite effect. We validate our theoretical findings on synthetic and real-world datasets. Our results are preliminary in the sense that, for now, we study only the case of one sensitive attribute, and only statistical disparity, conditional statistical disparity, and equal opportunity difference.

Related publication: [17]

8.1.8 Entrepôts de Données de Santé et Protection de la Vie Privée: Synthèse de discussions Inter-CHU

Participants: Antoine Boutet, et al..

Ces dernières années, la mise en chantier de différents EDS a fait émerger des discussions, entre divers acteurs travaillant sur ces EDS, concernant la protection de la vie privée des patients. Cet article présente une synthèse des points abordés durant ces discussions. Nous y argumentons que les définitions législatives offrent une base de travail solide. Nous concluons que la multiplicité et la mixité des méthodes offrent la meilleure protection de la vie privée pour les patients, même si celles-ci doivent s'adapter en fonction des besoins des investigateurs.

English version: During the last years, the construction of Health Data Warehouses in France raised discussions about Patient Privacy. This article proposes a synthesis of the points covered during these

discussions. We argue that legal definitions offer a strong base of work. We conclude that the diversity and the multiplicity of the methods used allow the best Patient Privacy, even if the use of these methods must be adapted according to the needs of investigators.

Related publication: [27]

8.1.9 A Smartphone-based Architecture for Prolonged Monitoring of Gait

Participants: Mohamed Bechorfa, Antoine Boutet, et al..

Gait analysis is important for evaluating neurological disorders such as stroke and Parkinson's disease. Traditionally, healthcare professionals had to rely on subjective assessments (i.e., human-based) of gait which were time consuming and not very reproducible. However, with the advent of IoT and indeed more objective (e.g., measurement-based) assessment methods, gait analysis can now be performed more accurately and effectively. It is worth noting, however, that there are still limitations to these objective methods, especially the lack of privacy-preserving continuous data collection. To overcome this limitation, we present in this worl a privacy-by-design monitoring application for post-stroke patients to evaluate their gait before, during, and after a rehabilitation program. Gait measurements are collected by a mobile application that continuously captures spatiotemporal parameters in the background using the built-in smartphone accelerometer. Statistical techniques are then applied to extract general indicators about the performed activity, as well as some more specific gait metrics in real-time such as regularity, symmetry and walking speed. These metrics are calculated based on the detected steps while patients are performing an activity. Additionally, a deep learning approach based on an auto-encoder is implemented to detect abnormal activities in the gait of patients. These analyses provides both valuable insights and statistical information about the activities performed by the patient, and a useful tool for practitioners to monitor the progression of neurological disorders and detect anomalies. We conducted experiments using this application in real conditions to monitor post-stroke patients in collaboration with a hospital, demonstrating its ability to compute valuable metrics and detect abnormal events patient's gait.

Related publication: [28]

8.1.10 Causal Discovery Under Local Privacy

Participants: Héber Arcolezi, et al..

Differential privacy is a widely adopted framework designed to safeguard the sensitive information of data providers within a data set. It is based on the application of controlled noise at the interface between the server that stores and processes the data, and the data consumers. Local differential privacy is a variant that allows data providers to apply the privatization mechanism themselves on their data individually. Therefore, it provides protection also in contexts in which the server, or even the data collector, cannot be trusted. The introduction of noise, however, inevitably affects the utility of the data, particularly by distorting the correlations between individual data components. This distortion can prove detrimental to tasks such as causal structure learning. In this paper, we consider various well-known locally differentially private mechanisms and compare the trade-off between the privacy they provide, and the accuracy of the causal structure produced by algorithms for causal learning when applied to data obfuscated by these mechanisms. Our analysis yields valuable insights for selecting appropriate local differentially private protocols for causal discovery tasks. We foresee that our findings will aid researchers and practitioners in conducting locally private causal discovery.

Related publication: [13]

8.2 Research axis 2: Web, smartphone, IoT, and wireless

8.2.1 Web, smartphone, AdTech: the privacy viewpoint

Participants: Vincent Roca, et al..

This tutorial addresses the topic of personal data collection in web and smartphone environments, the management of user consent, and the underlying ecosystem setup by the AdTech industry to build user profiles and use them in real-time bidding (RTB) systems. This tutorial does not enters complex technical details. Its goal is to provide an overview and identify trends, with on the one hand, a highly profitable business that benefits from new data sources and advanced profiling technologies, and on the other hand, a business that is more and more limited by strict regulations and a growing privacy and sustainability awareness.

Related publication: [10]

8.2.2 RSSI-based attacks for identification of BLE devices

Participants: Mathieu Cunche, et al..

To prevent tracking, the Bluetooth Low Energy (BLE) protocol integrates privacy mechanisms such as address randomization. However, as highlighted by previous researches address randomization is not a silver bullet and can be circumvented by exploiting other types of information disclosed by the protocol such as counters or timing. In this work, we propose two novel attack to break address randomization in BLE exploiting side information in the form of Received Signal Strength Indication (RSSI). More precisely, we demonstrate how RSSI measurements, extracted from received BLE advertising packets, can be used to link together the traces emitted by the same device or directly re-identify it despite address randomization. The proposed attacks leverage the distribution of RSSI to create a fingerprint of devices with an empirical evaluation on various scenarios demonstrating their effectiveness. For instance in the static context, in which devices remain at the same position, the proposed approach yields a re-identification accuracy of up to 97%, which can even be boosted to perfect accuracy by increasing the number of receivers controlled by the adversary. We also discuss the factors influencing the success of the attacks and evaluate two possible countermeasures whose effectiveness is limited, highlighting the difficulty in mitigating this threat.

Related publication: [6]

8.2.3 Enhancing IoT Privacy: Why DNS-over-HTTPS Alone Falls Short?

Participants: Samuel Pelissier, Abhishek Kumar Mishra, Mathieu Cunche, et al..

Recent years have seen widespread adoption of consumer Internet of Things (IoT) devices, offering diverse benefits to end-users, from smart homes to healthcare monitoring, but raising serious privacy concerns. To address this, securing efforts, such as encrypting DNS, have been proposed. In this paper, we study the effectiveness of such measures in the specific context of ensuring IoT privacy. We introduce a device identification attack against DNS-over-HTTPS-enabled IoT devices. We conduct more than 25,000 automated experiments across 6 public DNS resolvers and find that the proposed attack can identify devices via DNS-over-HTTPS (DoH) traffic with a 0.98 balanced accuracy. We point out padding as a mitigation technique that reduces identification by a significant 33%. Additionally, we find that half of the evaluated DNS resolvers do not adhere to the relevant specification, substantially compromising user privacy.

Related publication: [24]

8.2.4 Fingerprinting Connected Wi-Fi Devices Using Per-network MAC Addresses

Participants: Abhishek Kumar Mishra, Samuel Pelissier, Mathieu Cunche.

Wi-Fi stands out as one of the most prominent and widespread wireless technologies in use today. Smartphones and various other Wi-Fi-enabled devices employ management frames called probe-requests to discover nearby networks. In this study, we reveal that it is possible to fingerprint based on the probe-requests they emit while connected to a network. Leveraging distinctive features of probe-request bursts we use a Random Forest-based approach to successfully fingerprint devices. This demonstrate that device randomizing their MAC addresses between networks can still be tracked. Through an assessment conducted on a real-world measurement comprising Wi-Fi devices with diverse operating systems, and spanning a month duration, we demonstrate that our model fingerprints individual devices with around 40% accuracy with 1 burst and perfect re-identification if two or more bursts are available.

Related publication: [21]

8.2.5 Third eye: Inferring the State of Your Smartphone Through Wi-Fi

Participants: Abhishek Kumar Mishra, Mathieu Cunche.

Wi-Fi is one of the most notable and prevalent wireless technologies today. Smartphones and other Wi-Fi-enabled devices find nearby networks using management frames known as probe-requests. In this paper, we infer the state of smartphones by passively monitoring their transmitted probe-requests. We leverage the differential behaviour of probe-request bursts and their content, based on their device states such as active/static screen and Wi-Fi/power-saving mode ON/OFF. We use a Random Forest based approach that can successfully predict smartphone states just leveraging individual bursts. Based on an evaluation using a real-world dataset of more than 200 smartphones (having a variety of operating systems), with ground truth data available, we show that our model reliably predicts states with accuracy >=98%.

Related publication: [19]

8.2.6 Revealing Disparities in Public and Digital Infrastructure of Developing Countries

Participants: Abhishek Kumar Mishra, et al..

This work, for the first time in the literature, investigates and establishes the disparities between public infrastructure and digital inequality by employing NetMob 2024 datasets. We first define the inference of the state of these infrastructures before quantifying infrastructure inequality using the Gini coefficient in developing countries, i.e., Colombia, India, Indonesia, and Mexico, and unveil the synergy between public and digital infrastructure by leveraging Spearman's rank correlation. The categorization of origin-destination NetMob 2024 data into half-yearly chunks uncovers the varying evolution of the public and digital infrastructure. The results reveal the existence of substantial disparities and temporal variations between the recorded year's halves.

Related publication: [20]

8.2.7 Post-Quantum Authentication and Integrity in 3-Layer IoT Architectures

Participants: Clémentine Gritti, et al..

The Internet of Things (IoT) is a growing area of technology and has been identified as a key tool for enhancing industries' operation and performance. As IoT deployment rises worldwide, so do the

threats; hence, security, especially authentication and integrity, is a critical consideration. One significant future threat is quantum attacks, which can only be defeated using Post-Quantum (PQ) cryptosystems. New Digital Signature (DS) standards for PQ security have been selected by the US National Institute of Standards and Technology (NIST). However, IoT comes with its own technical challenges from the constrained resources allocated to sensors and other similar devices. As a consequence, the use and suitability of these PQ schemes for IoT remains an open research area. In this paper, we identify an IoT architecture built from three distinct layers represented by a server, a gateway and an IoT device, respectively. We first test PQ DS scheme standards and compare them with current standards in order to assess their practicality for use in this architecture to provide authentication and integrity. Then, we select the most suitable PQ scheme at each layer according to the features of the corresponding device (server, gateway, IoT device) and the security property (authentication, integrity). We finally carry out experiments on our selection and provide an architectural model for making IoT communication and interaction PQ secure.

Related publication: [25]

8.2.8 MATRaCAE: Time-based Revocable Access Control in the IoT

Participants: Clémentine Gritti, et al..

Internet of Things (IoT) promises a strong connection between digital and physical environments. Nevertheless, this framework comes with security vulnerabilities, due to the heterogeneous nature of devices and the diversity of their provenance. Furthermore, technical constraints (e.g. devices' limited resources) require to lighten the design of the underlying security protocols. Liu et al. presented a system for data access with time-based control and direct user revocation that are beneficial features in IoT. In this paper, we propose an extension of this system, called MATRaCAE, that involves multiple authorities and considers binary time credentials. Doing so, we mitigate the key escrow problem and comes with a better trade-off between key update frequency and number of revoked users, which limited the applicability of Liu et al.'s scheme in IoT. Our solution can be proved secure under the Decisional Bilinear Diffie-Hellman Exponent assumption. Subsequently, we implement and evaluate MATRaCAE to demonstrate its suitability to IoT frameworks.

Related publication: [15]

8.2.9 Characterizing probe request bursts to efficiently count Wi-Fi devices with randomized MACs

Participants: Abhishek Kumar Mishra, Mathieu Cunche.

In this paper, we show that counting the number of devices in a geographical zone is possible by passively capturing Wi-Fi probe-requests, even in the presence of randomized MAC addresses. We utilize a clustering-based approach and carefully characterize probe-request bursts with features that tend to be specific to a device. On three datasets with different capture configurations, we show that our methodology successfully counts the number of devices with a maximum error of 1 device.

Related publication: [18]

8.2.10 Privacy-Preserving Pseudonyms for LoRaWAN

Participants: Samuel Pélissier, Jan Aalmoes, Abhishek Kumar Mishra, Mathieu Cunche, Vincent Roca, et al..

LoRaWAN, a widely deployed LPWAN protocol, raises privacy concerns due to metadata exposure, particularly concerning the exploitation of stable device identifiers. For the first time in literature, we

propose two privacy-preserving pseudonym schemes tailored for LoRaWAN: resolvable pseudonyms and sequential pseudonyms. We extensively evaluate their performance and applicability through theoretical analysis and simulations based on a large-scale real-world dataset of 71 million messages. We conclude that sequential pseudonyms are the best solution.

Related publication: [23]

8.2.11 Collecte de traces WiFi publiques: de la protection de la vie privée à l'analyse de trajectoires

Participants: Abhishek Kumar Mishra, et al..

Dans le paysage actuel marqué par l'omniprésence des smartphones et des réseaux sans fil, la génération d'empreintes numériques est devenue courante, révélant les habitudes quotidiennes des utilisateurs. Cet article présente un ensemble d'outils pour collecter et analyser des traces WiFi. Ces outils relèvent différents défis, tels que la gestion des associations des adresses MAC des smartphones, la reconstruction de trajectoires des utilisateurs et la protection de leur confidentialité. En abordant systématiquement ces défis, ces outils visent à faciliter la compréhension de la mobilité des individus et à établir des contacts plausibles entre divers appareils.

English version: In today's landscape marked by the ubiquity of smartphones and wireless networks, the generation of digital footprints has become commonplace, revealing users' daily habits. This article presents a set of tools for collecting and analyzing WiFi traces. These tools address various challenges, such as managing smartphone MAC address associations, reconstructing user trajectories and protecting user confidentiality. By systematically tackling these challenges, these tools aim to facilitate understanding of people's mobility and establish plausible contacts between various devices.

Related publication: [22]

8.3 Research axis 3: User empowerment

8.3.1 The Effect of Design Patterns on (Present and Future) Cookie Consent Decisions

Participants: Nataliia Bielova, et al..

Today most websites in the EU present users with a consent banner asking about the use of cookies or other tracking tech- nologies. Data Protection Authorities (DPAs) need to ensure that users can express their true preferences when faced with these banners, while simultaneously satisfying the EU GDPR requirements. To address the needs of the French DPA, we conducted an online experiment among 3,947 participants in France exploring the impact of six different consent banner designs on the outcome of users' consent decision. We also assessed participants' knowledge and privacy preferences, as well as satisfaction with the banners. In contrast with previous results, we found that a "bright pattern" that highlights the decline option has a substantial effect on users' decisions. We also find that two new designs based on behavioral levers have the strongest effect on the outcome of the consent decision, and participants' satisfaction with the banners. Finally, our study provides novel evidence that the effect of design persists in a short time frame: designs can significantly affect users' future choices, even when faced with neutral banners.

Related publication: [12]

8.3.2 An Ontology of Dark Patterns: Foundations, Definitions, and a Structure for Transdisciplinary Action

Participants: Nataliia Bielova, Cristiana Santos, et al..

Deceptive and coercive design practices are increasingly used by companies to extract profit, harvest data, and limit consumer choice. Dark patterns represent the most common contemporary amalgamation of these problematic practices, connecting designers, technologists, scholars, regulators, and legal professionals in transdisciplinary dialogue. However, a lack of universally accepted definitions across the academic, legislative, practitioner, and regulatory space has likely limited the impact that scholarship on dark patterns might have in supporting sanctions and evolved design practices. In this paper, we seek to support the development of a shared language of dark patterns, harmonizing ten existing regulatory and academic taxonomies of dark patterns and proposing a three-level ontology with standardized definitions for 64 synthesized dark pattern types across low-, meso-, and high-level patterns. We illustrate how this ontology can support translational research and regulatory action, including transdisciplinary pathways to extend our initial types through new empirical work across application and technology domains.

Related publication: https://doi.org/10.1145/3613904.3642436

8.3.3 Two worlds apart! Closing the gap between regulating EU consent and user studies

Participants: Nataliia Bielova, Cristiana Santos, et al..

The EU ePrivacy Directive requires consent before using cookies or other tracking technologies, while the EU General Data Protection Regulation ("GDPR") sets high-level and principle-based requirements for such consent to be valid. However, the translation of such require- ments into concrete design interfaces for consent banners is far from straightforward. This situation has given rise to the use of manipulative tactics in user experience ("UX"), commonly known as dark patterns, which influence users' decision-making and may violate the GDPR re- quirements for valid consent. To address this problem, EU regulators aim to interpret GDPR requirements and to limit the design space of consent banners within their guidelines. Academic researchers from various disciplines address the same problem by performing user stud- ies to evaluate the impact of design and dark patterns on users' decision making. Regrettably, the guidelines and user studies rarely impact each other. In this Essay, we collected and analyzed seventeen official guide- lines issued by EU regulators and the EU Data Protection Board ("EDPB"), as well as eleven consent-focused empirical user studies which we thoroughly studied from a User Interface ("UI") design per- spective. We identified numerous gaps between consent banner designs recommended by regulators and those evaluated in user studies. By do- ing so, we contribute to both the regulatory discourse and future user studies. We pinpoint EU regulatory inconsistencies and provide action- able recommendations for regulators. For academic scholars, we syn- thesize insights on design elements discussed by regulators requiring further user study evaluations. Finally, we recommend that EDPB and EU regulators, alongside usability, Human-Computer Interaction ("HCI"), and design researchers, engage in transdisciplinary dialogue in order to close the gap between EU guidelines and user studies.

Related publication: [3]

8.4 Research axis 4: Legal

8.4.1 The lawfulness of re-identification under data protection law

Participants: Teodora Curelariu, Alexandre Lodie.

Data re-identification methods are becoming increasingly sophisticated and can lead to disastrous data breaches. Re-identification is a key research topic for computer scientists as it can be used to reveal vulnerabilities of de-identification methods such as anonymisation or pseudonymisation. However, re-identification, even for research purposes, involves processing personal data. From this background, this paper aims to investigate whether reidentification carried out by computer scientists for research purposes can be considered GDPR-compliant. This issue is paramount to contribute to improving the state of knowledge concerning data security measures.

Related publication: [9]

8.4.2 Is it Personal data? Solving the gordian knot of anonymisation

Participants: Alexandre Lodie, Cédric Lauradoux.

The concept of personal data is pivotal to understand the scope of the General Data Protection Regulation (GDPR). Since data protection regulations and directives were adopted, national courts and the Court of Justice of the European Union (CJEU) had to determine whether some data like IP addresses are personal. Courts' rulings are often based on the possibility to re-identify individuals from the datasets under dispute. The different views adopted by Courts over the years do not always reach the same conclusions, which is source of legal uncertainties. This is especially the case when data controllers are using data protection techniques like pseudonymisation and anonymisation. Recently, the ruling of the CJEU in the SRB vs EDPS case challenged the stance adopted by data protection authorities concerning the distinction between pseudonymisation and anonymisation. Data protection watchdogs consider that pseudonymized data are always personal data. The dictum of the court in the SRB vs EDPS case is that pseudonymised data can be considered as anonymised and thus non-personal data depending on the re-identification capability of the data holder. This creates legal uncertainties as the legal qualification of data that have been subject to data protection techniques. In this paper, the authors question the extent of the definition of personal data and how it applies to data protection techniques such as pseudonymisation and anonymisation. Eventually, they emphasise that this issue is challenging with regard to the protection of data within the EU borders and beyond.

Related publication: [29]

8.4.3 Standardised Messenger Audit

Participants: Mathieu Cunche, et al..

The Standardised Messenger Audit project aims at helping to inspect the messenger services used within businesses from a data protection perspective. The EDPB launched the Standardised Messenger Audit project in the context of the Support Pool of Experts programme at the request of the German Federal Data Protection Authority (DPA).

Related publication: [33]

8.4.4 Feedback to the European Data Protection Board's Guidelines 2/2023 on Technical Scope of Art. 5(3) of ePrivacy Directive

Participants: Cristiana Santos, Nataliia Bielova, Vincent Roca, Mathieu Cunche, Gilles Mertens, Karel Kubicek, et al..

This project aims to give feedback to the European Data Protection Board (EDPB) on its Guidelines 2/2023 on Technical Scope of Art. 5(3) of ePrivacy Directive published by the EDPB in the end of 2023.

Related publication: [34]

9 Bilateral contracts and grants with industry

9.1 Bilateral contracts with industry

OpenSezam

Participants: Mohamed Maouche, Vincent Roca.

- CIFRE PhD contract, 2024-2026 (OpenSezam website)
- The thesis is entitled: *"Secure, Private and Multi-modal Authentication"*. The objective is to design an authentication system based on end-to-end machine learning, with an integrated system for continuous detection of anomalies and intrusions, using various types of biometric data, depending on the use-case.

Murena

Participants: Vincent Roca, Mathieu Cunche.

- CIFRE PhD contract, 2024-2026 (Murena website)
- The thesis is entitled: *"Blocking trackers via Federated ML and integration in the /e/OS smartphone operating system"*. It aims to integrate a system for identifying and blocking undesirable flows into the /e/OS operating system, a mobile operating system privacy-friendly by construction, rid of any Google components. This work will focus on two main directions. Firstly, the development of an automated blocking approach federating a population of users. Secondly, the integration of this solution into the e/OS/ operating system.

10 Partnerships and cooperations

10.1 International initiatives

10.1.1 Inria associate team not involved in an IIL or an international program

AUDIT-PAIR

Title: Algorithmic Auditing of Privacy and Fairness

Duration: 2024 - 2026

Coordinator: Sébastien Gambs (gambs.sebastien@uqam.ca)

Partners:

- Université du Québec À Montréal (Canada)
- ÉTS Montréal (Canada)
- University of Ottawa (Canada)

Inria contact: Heber Hwang Arcolezi

Summary: The AUDIT-PAIR Associate Team is a collaborative research initiative involving Privatics and three Canadian partners (UQAM, ÉTS Montréal, and the University of Ottawa). Our goal is to advance research and development in AI system auditing, with a specific focus on privacy and fairness. In a world increasingly reliant on algorithmic decision-making systems, we recognize the critical challenges and ethical responsibilities that accompany these technologies. AUDIT-PAIR is dedicated to addressing the complex challenges related to privacy and fairness audits in AI systems. We aim to propose practical methodologies and systems to tackle these issues effectively. We expect AUDIT-PAIR to play a pivotal role in promoting the adoption of privacy and fairness auditing as a fundamental and standardized practice in the development, deployment, and maintenance of AI systems.

MAGPIE

Title: Machine Learning and privacy challenges

Duration: 2022 -> 2024

Coordinator: Emiliano De Cristofaro (e.decristofaro@ucl.ac.uk)

Partners:

- University College London London (Royaume-Uni)

Inria contact: Mathieu Cunche

Summary: Machine learning offers great potential but also comes with a number of drawbacks. In particular when ML is applied to personal data, the privacy of individual may be at risks. In an orthogonal approach, ML can be leveraged to tackle privacy issues, for instance by sanitizing data or automatically detecting issues in complex systems. In MAGPIE, two teams from UCL and Inria with a strong expertise in privacy and machine learning will collaborate to explore those two research directions.

10.2 International research visitors

10.2.1 Visits of international scientists

Inria International Chair

Participants: Cristiana Teixeira Santos.

Cristiana is Lecturer and Researcher in Law and Technologies at Utrecht University, NL. She was awarded an Inria International Chair (2023-2025) and regularly visits Inria Sophia Antipolis as SRP researcher.

Other international visits to the team

Juliet Samandari

Status: PhD student

Institution of origin: University of Canterbury

Country: New Zealand

Dates: Oct-Nov 2024

Context of the visit: co-supervision of the thesis

Mobility program/type of mobility: research visit

Kumari Kancherla

Status PhD student

Institution of origin: IIT Gandhinagar

Country: India

Dates: Apr-Jun 2024

Context of the visit: collaboration on a research project

Mobility program/type of mobility: research visit

10.2.2 Visits to international teams

Research stays abroad

Héber Hwang Arcolezi

Status: ISFP

Institution visited: Université du Québec à Montréal (UQAM)

Country: Canada

Dates: Aug-Sept 2024 (3 weeks)

Context of the visit: Collaboration within the AUDIT-PAIR associated team. Invited Speaker at (SAC 2024 Summer School.

Héber Hwang Arcolezi

Status: ISFP

Institution visited: Google NY

Country: USA

Dates: Aug 2024 (3 days)

Context of the visit: Discussion with Google Privacy researchers on potential topics for future collaboration.

10.3 National initiatives

10.3.1 ANR

IPoP (PEPR Cybersecurity)

- Title: Interdisciplinary Project on Privacy
- Type: PEPR Cybersécurité / France 2030
- Duration: July 2022 - June 2028
- Coordinator: Inria - PRIVATICS
- Others partners: Inria COMET / MAGNET / PETRUS / MULTISPEECH and SPIRALS teams, CNRS - DCS lab., INSA CVL - LIFO lab., Univ. Grenoble Alps - CESICE lab., Univ. of Rennes 1 - SPICY team, EDHEC, CNIL
- Abstract: IPoP focuses on new forms of personal information collection, on 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 recognized research teams (universities, engineering schools and institutions) and the CNIL.

SSF-ML-DH (PEPR Santé Numérique)

- Title: Secure, safe and fair machine learning for healthcare
- Type: PEPR Santé Numérique / France 2030
- Duration: November 2023 - October 2027
- Coordinator: Inria - MAGNET
- Others partners: Inria PRIVATICS, Inria EPIONE, CNRS Lamsade, Cea - List, IMT Atlantique, CNRS Diens
- Abstract: The healthcare sector generates vast amounts of data from various sources (e.g., electronic health records, imaging, wearable devices, or population health data). These datasets, analyzed through ML systems, could improve the whole healthcare system, for the individuals and the society. However, the sensitive nature of health data, cybersecurity risks, biases in the data, and the lack of robustness of ML algorithms are all factors that currently limit the widespread use of such data bases. The project aims to develop new ML algorithms, designed to handle the unique characteristics of multi-scale and heterogeneous individual health data, while providing formal privacy guarantees, robustness against adversarial attacks and changes in data dynamics, and fairness for under-represented populations.

TULIP

- Title: TULIP
- Type: ANR MRSEI (Montage de Réseaux Scientifiques Européens ou Internationaux)
- Duration: 2024 - 2025
- Coordinator: Inria - PRIVATICS, Nataliia Bielova
- Others partners: University of Utrecht
- Abstract: The TULIP (proTéger les UtiLisateurs contre les manIPulations en ligne) project funded by the ANR MRSEI program. In order to protect users from manipulation, deception, and harms caused by dark patterns in digital systems, the TULIP project aims at advancing the knowledge about the cognitive mechanisms used by dark patterns, developing new methods for automatic detection of dark patterns across contexts and tools to help regulators collect evidence of dark patterns. To achieve these ambitious goals, in this project we will advance the research in dark patterns from three dimensions: legal, human-computer interaction and computer science. This project aims at building the consortium to respond to the ERC Synergy program and covers costs to meet with the other co-PIs for the synergy grant and advance the project proposal development.

Frugal Internet

- Title: Development of a Frugal Internet - Networks and Systems with Reduced and Sustainable Energy and Carbon Constraints
- Type: ANR grant associated to Clémentine Gritti's Chaire Professeur Junior position
- Duration: 2024 - 2029
- Coordinator: Inria - PRIVATICS, Clémentine Gritti
- Abstract: Two concurrent developments are shaping the future of the Internet: (1) driven by the pandemic and remote work, a significant digital transition with growing demands for efficient communication and computing services, and (2) under the climate imperative, a shift towards a digital transition that is energy-efficient, low-carbon, and sustainable. These two seemingly contradictory developments pave the way for a more reasoned Internet. Currently, each component

of the Internet is deployed independently, achieving localized energy efficiency, but must be reimagined to create globally frugal networks and systems "by design." The key point here is to establish a fully optimized information processing chain as close to the user as possible—globally optimized while remaining autonomous, sovereign, and deployable by a local operator.

PMR

- Title: Privacy-preserving methods for Medical Research
- Type: ANR
- Duration: 2020 - 2024
- Coordinator: Inria MAGNET
- Others partners: INSA Lyon, Creatis
- Abstract: Given the growing awareness of privacy risks of data processing, there is an increasing interest in privacy-preserving learning. However, shortcomings in the state of the art limit the applicability of the privacy-preserving learning paradigm. First, most approaches assume too optimistically a honest-but-curious setting. Second, most approaches consider one learning task in isolation, not accounting for the context where querying is a recurring activity. In this project, we will investigate new algorithms and models that address these shortcomings. Among others, (i) our algorithms will combine privacy-preserving properties of differential privacy with security offered by cryptography and (ii) based on models of information flows in integrated data handling processes, we will build more refined models analyzing the implications of repeated querying. We will demonstrate the utility of our new theory and algorithms by proposing strategies to realistically apply them in significant real-world problems illustrated through use cases in the medical domain.

10.3.2 Inria Exploratory Action (AEx)

DATA4US (Personal DAta TrAnsparency for web USers)

- Participants: Cedric Lauradoux, Nataliia Bielova
- Duration: 2020-2024
- Abstract: Since May 2018, General Data Protection Regulation (GDPR) regulates collection of personal data in all EU countries, but users today are still tracked and their data is still silently collected as they browse the Web. GDPR empowers users with the rights to access their own data, but users have no means to exercise their rights in practice. DATA4US tackles these interdisciplinary challenges by establishing collaborations with researchers in Law. DATA4US will propose a new architecture for exercising access rights that will explain the users whether their data has been legally collected and eventually help contact DPAs for further investigations.

10.4 Regional initiatives

INTERFERE

- Title: Stressing Systems Security Through on the Fly Network Traffic Generation
- Type: FIL (Federation Informatique Lyonnaise)
- Duration: 2024 - 2025
- Coordinator: ENS/LIP
- Others partners: Inria PRIVATICS

- Abstract: The INTERFERE project focuses on the generation of sequential data such as network traffic or query flows to stress systems against security issues. In particular, we consider intrusion attempts or attempts to re-identify users through multiple queries on an anonymous data warehouse. The INTERFERE project aims to generate large quantities of data on the fly with a control of the events represented. By investigating audit tools, this project plays a key role in the development of new methods and technologies to strengthen the security of systems and data.

10.5 Public policy support

10.5.1 INRIA-CNIL collaboration

PRIVATICS and CNIL collaborate since 2012 through several collaborative projects:

- the Mobilitics bi-lateral project on privacy and smartphones in 2012-2014;
- the IoTics ANR research project on privacy and connected devices;
- since 2022, CNIL is part of the consortium of the IPoP project (PEPR cybersecurity).

Several workshops and discussions on data anoymisation, risk analysis, consent or IoT Privacy also took place over the years, and more recently, we contributed (e.g., with other IPoP partners) to the reviewing of candidate CNIL position documents for AI in the context of public consultations.

PRIVATICS is also in charged of the organization of the CNIL-Inria prize that is awarded every year to an outstanding publication in the field of data privacy.

Last but not least:

- Nataliia Bielova worked in the Laboratoire d'Innovation de la CNIL (LINC) in the context of a "mise à disposition", from September 2021 to December 2022. News CNIL
- On August 2021, Claude Castelluccia was appointed member of the CNIL Board ("commissaire CNIL") as one qualified public figures for his expertise on digital sciences and privacy questions. As such he is in charge of several domains and contributes to the doctrine of the French Data Protection Authority. This is a *major involvement representing approximately 80% of his professional activity*.

11 Dissemination

11.1 Promoting scientific activities

11.1.1 Scientific events: organisation

- APVP (Atelier sur la Protection de la Vie Privée) 2024, Juin 2024, le domaine Lou Capitelle (Mohamed Maouche, Héber H. Arcolezi, Mathieu Cunche, Antoine Boutet)
- Atelier IPoP : Audits de l'IA, Mai 2024, Paris (Antoine Boutet)
- Mobilizing Research and Regulatory Action on Dark Patterns and Deceptive Design Practices workshop at the ACM CHI conference aiming to unite transdisciplinary researchers in law, computer science and design, May 2024, Honolulu (Nataliia Bielova)
- Privacy Alpine Seminar, March 2024, Corrençon en Vercors (Mathieu Cunche, Antoine Boutet)
- GDR RSD / ASF Winter School on Distributed Systems and Networks, February 2024, Le Pleynet (Antoine Boutet)
- Annual Privacy Forum (APF) 2024, September 2024, Karlstad, Sweden (Cedric Lauradoux)

Member of the organizing committees

- Matinée PTCC : Cybersécurité en santé - allier sécurité et partage (Antoine Boutet)

11.1.2 Scientific events: selection

Member of the conference program committees

- ACM WiSec 2024-2025, Mathieu Cunche
- APF 2024, Cédric Lauradoux
- IFIPSC 2024, Cédric Lauradoux
- DPM 2024, Mathieu Cunche
- SINCONF 2024, Clementine Gritti
- CCS 2024, Héber H. Arcolezi
- PETS 2024, Héber H. Arcolezi
- IJCAI 2024, Héber H. Arcolezi
- FAccT 2024, Héber H. Arcolezi
- SAC 2024, Héber H. Arcolezi
- ICLR 2024, Héber H. Arcolezi
- European Privacy Law Scholars Conference - Europe 2024 (PLSC), Nataliia Bielova

Reviewer Too many to be listed.

11.1.3 Journal

Member of the editorial boards

- IACR Communications in Cryptology, Clementine Gritti

Reviewer - reviewing activities

- JISA, Mathieu Cunche
- The Lancet, Antoine Boutet
- Springer Data Mining and Knowledge Discovery, Antoine Boutet
- International Journal of Information Security, Nataliia Bielova
- IEEE Sensors Journal, Vincent Roca
- IEEE Transactions on Information Forensics and Security, Vincent Roca

11.1.4 Invited talks

- December 2024: Antoine Boutet - PEPR Cyber Day, Paris – "Données de santé : allier confidentialité et partage"
- December 2024: Antoine Boutet - Les journées scientifiques du PEPR Cybersécurité
- October 2024: Claude Castelluccia - Hybrid-Mind conference à Genève,"Cognitive Security"
- Novembre 2024: Antoine Boutet - Matinée Cyber et Santé. Programme de Transfert au Campus Cybersécurité
- October 2024: Nataliia Bielova - CNIL's Seminar on "Cookies and Web Tracking".

- October 2024: Nataliia Bielova - "Digital Regulation: the contribution of science" event organised by Inria in Brussels, Belgium.
- August 2024: Héber H. Arcolezi - Securing Data with Local Differential Privacy: Concepts, Protocols, and Practical Applications. Selected Areas in Cryptography Summer School 2024
- July 2024: Nataliia Bielova - *keynote* at the Privacy Enhancing Technologies Symposium (PETS), Bristol, UK.
- June 2024: Antoine Boutet - Webinaire BERNOULLI LAB – "L'anonymisation des données de santé est-elle possible" ?
- June 2024: Nataliia Bielova - *keynote* at the Mapping and Governing the Online Worls (MGOW), interdisciplinary conference bringing researchers from law, economics, computer science and related fields. Ascona, Switzerland.
- May 2024: Nataliia Bielova - invited panel member at the CPDP, Brussels.
- May 2024: Nataliia Bielova - Towards Neuro Ethics rights Workshop, a series of workshops to tackle the questions of neurodata protection for law and computer science experts, Paris.
- April 2024: Nataliia Bielova - Norwegian Consumer Council (Forbrukerrdet) on resent research results in online tracking and dark patterns, Oslo, Noway.
- Februrary 2024: Antoine Boutet - Webinaire, "Entrepôt de Données de Santé et Confidentialité des Données"
- January 2024: Antoine Boutet - PEPR Cybersecurity Winter School, Autrans - "AI and Privacy, a CNIL perspective"
- January 2024: Vincent Roca and Pierre Laperdrix - PEPR Cybersecurity Winter School, Autrans - "Web, smartphone, AdTech: the privacy viewpoint", [10]
- January 2024: Nataliia Bielova - Lorentz center seminar on Fair patterns for online interfaces on impact of design on users consent decisions. Leiden, the Netherlands.

11.1.5 Scientific expertise

- Mathieu Cunche : creation of *Standardised Messenger Audit* [33], a methodology for evaluating the compliance of messenger application with regard to GDPR. Done as part of EDPB's Support Pool of Experts programme at the request of the German Federal Data Protection Authority (DPA).
- Nataliia Bielova: external reviewer for the Canadian Social Sciences and Humanities Research Council (SSHRC).
- Nataliia Bielova: member of a Full-time Assistant or Full Professor Position in Cybersecurity recruiting committee at École Polytechnique in 2024.
- Nataliia Bielova: expert for the EU Commission for the implementation of the EU Digital Services Act (DSA).

11.1.6 Research administration

- Mathieu Cunche is chair of the Privacy working group (GT-PVP) at GDR-Sécurité Informatique
- Antoine Boutet is treasurer of ACM Sigops France (ASF)

11.2 Teaching - Supervision - Juries

11.2.1 Teaching

Most of the PRIVATICS members' lectures are given at INSA-Lyon (Antoine Boutet, Mathieu Cunche and Clementine Gritti are associated professor at INSA-Lyon), at Grenoble Alps University (Claude Castelluccia, Vincent Roca and Cédric Lauradoux), and Université Côte d'Azur (Nataliia Bielova).

Most of the PRIVATICS members' lectures are on the foundations of computer science, security and privacy, as well as networking. The lectures are given to computer science students but also to business school students and to laws students.

Details of lectures:

- Master : Antoine Boutet, *Privacy*, 80h, INSA-Lyon, France.
- Master : Antoine Boutet, *Security*, 40h, INSA-Lyon, France.
- Undergraduate course : Mathieu Cunche, *Introduction to computer science*, 70h, L1, INSA-Lyon, France.
- Undergraduate course : Cédric Lauradoux, *Introduction to Python*, 70h, L1, University of Grenoble Alpes, France.
- Undergraduate course : Clementine Gritti, *Informatics and numerical society*, 32h, L1, INSA-Lyon, France.
- Master : Mathieu Cunche, *Wireless Security*, 6h, M2, INSA-Lyon, France.
- Master : Mathieu Cunche, *Privacy and Data protection*, 26h, M2, INSA-Lyon, France.
- Master : Mathieu Cunche, *Cryptography and Communication Security*, 18h, M1, INSA-Lyon, France.
- Master : Clementine Gritti, *Cryptography and Communication Security*, 16h, M1, INSA-Lyon, France.
- Master : Cédric Lauradoux, *Privacy and GDPR* , 12h, M2, University of Grenoble Alpes, France.
- Master : Claude Castelluccia, *Data Privacy*, 20h, M2, Skema, France.
- Master : Claude Castelluccia, *Cyber Security*, 6h, M2, Laws School of University of Grenoble Alpes, France.
- Master : Claude Castelluccia, *Data Privacy*, 6h, M2, Laws School of University of Grenoble Alpes, France.
- Master : Claude Castelluccia, *Data Privacy*, 12h, SKEMA Business School, Sophia-Antipolis, France.
- Master : Vincent Roca, *Wireless Communications*, 16h, M2, Polytech, University of Grenoble Alpes, France.
- Master : Vincent Roca, *Privacy in Smartphones*, 1.5h, M2 (University of Cote-d'Azur), France.
- Master : Vincent Roca, , 1.5h, M2 (University of Cote-d'Azur), France.

11.2.2 Juries

- Mathieu Cunche was a rapporteur for the PhD thesis of Myriam Clouet, université Paris-Saclay, 2024
- Mathieu Cunche was a rapporteur for the PhD thesis of Olivier Hureau, université Grenoble-Alpes, 2024
- Antoine Boutet was examinateur for the PhD thesis of Huiyu Li, université de Sophia Antipolis, November 2024

- Vincent Roca was a rapporteur for the HDR of Dr. Sami Zhioua, Institut Polytechnique de Paris, March 2024.
- Nataliia Bielova was a rapporteur for the PhD thesis of Nikhil Jha, Politecnico di Torino (Italy), February 2024.
- Nataliia Bielova was a rapporteur for the PhD thesis of Feiyang Tang, NTNU (Norway), April 2024.
- Nataliia Bielova was a rapporteur for the PhD thesis of Jean Luc Intumwayase, University of Lille (France), December 2024.

11.3 Popularization

11.3.1 Productions (articles, videos, podcasts, serious games, ...)

The team gets attached to engage in dialogue with society by explaining complex subjects related to privacy. Several articles have published in TheConversation newspaper or other more mainstream channels:

- Septembre 2024, Le Grand Moissonnage des Données, Le Monde Binaire.
- September 2024: IA : La guerre des données aura-t-elle lieu ? Usbek et Rica
- April 2024: L'AI Act, ou comment encadrer les systèmes d'IA en Europe [5] - TheConversation
- March 2024: Case C-479/22 P, Case C-604/22 and the limitation of the relative approach of the definition of "personal data" by the ECJ [36] - The European Law Blog
- February 2024: Protéger la vie privée des systèmes d'IA : l'ambition du projet iPoP - InCyber News
- January 2024: L'IA générative pourrait aussi servir à exploiter des données personnelles en toute sécurité : la piste des données synthétiques [4] - TheConversation

11.3.2 Participation in Live events

- Antoine Boutet - November 2024: Les Échappées inattendues du CNRS, Lyon

11.3.3 Others science outreach relevant activities

- Nataliia Bielova has transferred her research results on tracking, consent and impact of consent design on users' decision-making to the Dutch Data Protection Authority, January 2024.
- Nataliia Bielova together with Sanju Ahuja has transferred her research results on the ontology of dark patterns to the French regulators ARCOM, DGCCRF and the CNIL, December 2024.

12 Scientific production

12.1 Major publications

[1] N. Bielova, L. Litvine, A. Nguyen, M. Chammat, V. Toubiana and E. Hary. 'The Effect of Design Patterns on (Present and Future) Cookie Consent Decisions'. In: *USENIX Security Symposium (USENIX Security 24)*. 33rd USENIX Security Symposium (USENIX Security 24). Philadelphia (PA), United States: USENIX Association, 14th Aug. 2024, pp. 2813–2830. URL: https://hal.science/hal-04913806.

[2] H. H. Arcolezi and S. Gambs. 'Revealing the True Cost of Locally Differentially Private Protocols: An Auditing Perspective'. In: *Proceedings on Privacy Enhancing Technologies* 2024.4 (July 2024), pp. 123–141. DOI: 10.56553/popets-2024-0110. URL: https://inria.hal.science/hal-04644975.

12.2 Publications of the year

International journals

[3] N. Bielova, C. Santos and C. M. Gray. 'Two worlds apart! Closing the gap between regulating EU consent and user studies'. In: *Harvard Journal of Law & Technology* 37.3 (1st June 2024), pp. 1295–1333. URL: https://hal.science/hal-04913768 (cit. on p. 17).

[4] A. Boutet and A. Leautier. 'L'IA générative pourrait aussi servir à exploiter des données personnelles en toute sécurité : la piste des données synthétiques'. In: *The Conversation France* (24th Jan. 2024). URL: https://hal.science/hal-04901134 (cit. on p. 28).

[5] A. Boutet, J. Sénéchal, M. Bernelin and W. Letrone. 'L'AI Act, ou comment encadrer les systèmes d'IA en Europe'. In: *The Conversation France* (10th Apr. 2024). URL: https://hal.science/hal-04797429 (cit. on p. 28).

[6] G. Gagnon, S. Gambs and M. Cunche. 'RSSI-based attacks for identification of BLE devices'. In: *Computers & Security* (Aug. 2024), p. 104080. DOI: 10.1016/j.cose.2024.104080. URL: https://inria.hal.science/hal-04687696 (cit. on p. 13).

[7] H. H. Arcolezi and S. Gambs. 'Revealing the True Cost of Locally Differentially Private Protocols: An Auditing Perspective'. In: *Proceedings on Privacy Enhancing Technologies* 2024.4 (July 2024), pp. 123–141. DOI: 10.56553/popets-2024-0110. URL: https://inria.hal.science/hal-04644975 (cit. on p. 9).

[8] K. Makhlouf, H. Hwang Arcolezi, S. Zhioua, G. B. Brahim and C. Palamidessi. 'On the Impact of Multi-dimensional Local Differential Privacy on Fairness'. In: *Data Mining and Knowledge Discovery* (27th May 2024), pp. 1–24. DOI: 10.1007/s10618-024-01031-0. URL: https://hal.science/hal-04329938 (cit. on p. 9).

Invited conferences

[9] T. Curelariu and A. Lodie. 'The lawfulness of re-identification under data protection law'. In: *Privacy Technologies and Policy : 12th Annual Privacy Forum, APF 2024, Karlstad, Sweden, September 4–5, 2024, Proceedings*. Annual Privacy Forum 2024. Vol. 14831. Lecture Notes in Computer Science. Karlstad, Sweden: Springer, 1st Aug. 2024, pp. 112–131. DOI: 10.1007/978-3-031-68024-3_6. URL: https://hal.science/hal-04668779 (cit. on p. 18).

[10] V. Roca and P. Laperdrix. 'Web, smartphone, AdTech: the privacy viewpoint'. In: Winter school 2024 of the PEPR Cybersecurity. Autrans, France, 29th Jan. 2024. URL: https://inria.hal.science/hal-04550725 (cit. on pp. 13, 26).

International peer-reviewed conferences

[11] J. Aalmoes, V. Duddu and A. Boutet. 'On the Alignment of Group Fairness with Attribute Privacy'. In: WISE 2024 - 25th International Web Information Systems Engineering conference. Doha, Qatar, 2024, pp. 1–15. URL: https://hal.science/hal-04739862 (cit. on p. 9).

[12] N. Bielova, L. Litvine, A. Nguyen, M. Chammat, V. Toubiana and E. Hary. 'The Effect of Design Patterns on (Present and Future) Cookie Consent Decisions'. In: *USENIX Security Symposium (USENIX Security 24)*. 33rd USENIX Security Symposium (USENIX Security 24). Philadelphia (PA), United States: USENIX Association, 14th Aug. 2024, pp. 2813–2830. URL: https://hal.science/hal-04913806 (cit. on p. 16).

[13] R. Binkyte, C. Pinzón, S. Lestyán, K. Jung, H. Hwang Arcolezi and C. Palamidessi. 'Causal Discovery Under Local Privacy'. In: *Proceedings of Machine Learning Research*. Third Conference on Causal Learning and Reasoning. Vol. 236. Los Angeles, CA, United States, 9th May 2024, pp. 325–383. URL: https://hal.science/hal-04617032 (cit. on p. 12).

[14] C. Eichler, N. Champeil, N. Anciaux, A. Bensamoun, H. H. Arcolezi and J. M. de Fuentes. 'Nob-MIAs: Non-biased Membership Inference Attacks Assessment on Large Language Models with Ex-Post Dataset Construction'. In: WISE 2024 - 25th International Web Information Systems Engineering conference. Vol. 15438. Lecture Notes in Computer Science. Doha, Qatar: Springer Nature Singapore, 30th Nov. 2025, pp. 441–456. DOI: 10.1007/978-981-96-0570-5_32. URL: https://hal.science/hal-04670325 (cit. on p. 10).

[15] C. Gritti, E. Regnath and S. Steinhorst. 'MATRaCAE: Time-based Revocable Access Control in the IoT'. In: SECRYPT 2024 - 21st International Conference on Security and Cryptography. Dijon, France: SCITEPRESS - Science and Technology Publications, 15th Oct. 2024, pp. 274–285. DOI: 10.5220/0012825700003767. URL: https://hal.science/hal-04742788 (cit. on p. 15).

[16] T. Lebrun, L. Béziaud, T. Allard, A. Boutet, S. Gambs and M. Maouche. 'Synthetic Data: Generate Avatar Data on Demand'. In: The International Web Information Systems Engineering conference (WISE). Vol. 15440. Lecture Notes in Computer Science. Doha-Qatar, France: Springer Nature Singapore, 27th Nov. 2025, pp. 193–203. DOI: 10.1007/978-981-96-0576-7_15. URL: https://hal.science/hal-04715055 (cit. on p. 10).

[17] K. Makhlouf, T. Stefanović, H. H. Arcolezi and C. Palamidessi. 'A Systematic and Formal Study of the Impact of Local Differential Privacy on Fairness: Preliminary Results'. In: CSF 2024 - 37th IEEE Computer Security Foundations Symposium. Enschede, Netherlands: IEEE, 20th Sept. 2024, pp. 1–16. DOI: 10.1109/CSF61375.2024.00039. URL: https://inria.hal.science/hal-04832154 (cit. on p. 11).

[18] A. K. Mishra and M. Cunche. 'Characterizing probe request bursts to efficiently count Wi-Fi devices with randomized MACs'. In: EuCNC-6G Summit - 2024 European Conference on Networks and Communications & 6G Summit - Special Sessions. Antwerp, Belgium, 2024, pp. 1–3. URL: https://inria.hal.science/hal-04531452 (cit. on p. 15).

[19] A. K. Mishra and M. Cunche. 'Third eye: Inferring the State of Your Smartphone Through Wi-Fi'. In: *IEEE*. LCN 2024 - 49th IEEE Conference on Local Computer Networks. Caen, France, 2024, pp. 1–7. URL: https://hal.science/hal-04630691 (cit. on p. 14).

[20] A. K. Mishra and S. Mishra. 'Revealing Disparities in Public and Digital Infrastructure of Developing Countries'. In: *NetMob 2024 Book of Abstracts*. NetMob 2024. Washignton DC, United States, 2024, pp. 1–2. URL: https://hal.science/hal-04700746 (cit. on p. 14).

[21] A. K. Mishra, S. Pélissier and M. Cunche. 'Fingerprinting Connected Wi-Fi Devices Using Per-network MAC Addresses'. In: 17th International Symposium on Foundations & Practice of Security (FPS – 2024). Montreal, Canada, 9th Dec. 2024. URL: https://hal.science/hal-04655338 (cit. on p. 14).

[22] F. Molano Ortiz, A. K. Mishra, F. D. de M. Silva, A. Fladenmuller and L. H. M. K. Costa. 'Collecte de traces WiFi publiques: de la protection de la vie privée à l'analyse de trajectoires'. In: *CoRes 2024: 9èmes Rencontres Francophones sur la Conception de Protocoles, l'Évaluation de Performance et l'Expérimentation des Réseaux de Communication*. CoRes 2024 - 9èmes Rencontres Francophones sur la Conception de Protocoles, l'Évaluation de Performance et l'Expérimentation des Réseaux de Communication. Ed. by N. Achir and A. Carneiro Viana. Saint-Briac-sur-Mer, France, 2024, pp. 1–4. URL: https://hal.science/hal-04568193 (cit. on p. 16).

[23] S. Pélissier, J. Aalmoes, A. K. Mishra, M. Cunche, V. Roca and D. Donsez. 'Privacy-Preserving Pseudonyms for LoRaWAN'. In: 17th ACM Conference on Security and Privacy in Wireless and Mobile Networks (WiSec 2024). Seoul (Korea), France, 27th May 2024. DOI: 10.1145/3643833.3656120. URL: https://inria.hal.science/hal-04525080 (cit. on p. 16).

[24] S. Pélissier, G. Anselmi, A. K. Mishra, A. M. Mandalari and M. Cunche. 'Enhancing IoT Privacy: Why DNS-over-HTTPS Alone Falls Short?' In: TrustCom-2024 - 23rd IEEE International Conference on Trust, Security and Privacy in Computing and Communications. Sanya, China: IEEE, 2024, pp. 1–8. URL: https://inria.hal.science/hal-04777603 (cit. on p. 13).

[25] J. Samandari and C. Gritti. 'Post-Quantum Authentication and Integrity in 3-Layer IoT Architectures'. In: PST 2024 - 21st Annual International Conference on Privacy, Security, and Trust. Sydney, Australia, 5th Nov. 2024, pp. 1–26. URL: https://hal.science/hal-04766664 (cit. on p. 15).

[26] R. Taiello, M. Önen, C. Gritti and M. Lorenzi. 'Let Them Drop: Scalable and Efficient Federated Learning Solutions Agnostic to Stragglers'. In: *ACM Digital Library*. ARES 2024 - 19th International Conference on Availability, Reliability and Security. ARES '24: Proceedings of the 19th International Conference on Availability, Reliability and Security 13. Vienna, Austria: ACM, 30th July 2024, pp. 1–12. DOI: 10.1145/3664476.3664488. URL: https://hal.science/hal-04742784 (cit. on p. 11).

Conferences without proceedings

[27] M. Ahikki, M. Berard, C. Bouin, A. Boutet, S. Breant, A. Calliger, A. Cohen, J.-F. Couchot, D. Delamarre, C. Dunoyer, T. Fabacher, L. W. Gauthier, D. Gimbert, C. Girard-Chanudet, R. Griffier, M. Hilka, Y. Jacob, V. Jouhet, D. Laiymani, L. Moros, J. Muller, D. Pellecuer, T. Petit-Jean, A. Richard, M. Salaun, F. Talbot, P. Wajsburt and K. Yauy. 'Health Data Warehouses and Patient Privacy: Synthesis of Inter-Hospitals discussions'. In: Journée Santé et IA 2024. La Rochelle, France, 2024, pp. 1–5. URL: https://hal.science/hal-04613838 (cit. on p. 12).

[28] L. Bart, E. A. Bechorfa, A. Boutet, J. Ramon and C. Frindel. 'A Smartphone-based Architecture for Prolonged Monitoring of Gait'. In: 2024 IEEE First International Conference on Artificial Intelligence for Medicine, Health and Care (AIMHC). Laguna Hills, United States: IEEE, 5th Feb. 2024, pp. 16–17. DOI: 10.1109/AIMHC59811.2024.00010. URL: https://hal.science/hal-04909717 (cit. on p. 12).

[29] A. Lodie and C. Lauradoux. 'Is it Personal data? Solving the gordian knot of anonymisation'. In: Privacy Symposium 2024. Venise, Italy, 2024, pp. 1–18. URL: https://hal.science/hal-04609238 (cit. on p. 18).

Doctoral dissertations and habilitation theses

[30] A. Boutet. 'Privacy issues in AI and geolocation: from data protection to user awareness'. Insa Lyon, 10th Dec. 2024. URL: https://hal.science/tel-04909989 (cit. on p. 7).

[31] T. Lebrun. 'Health Data: Exploring Emerging Privacy Enhancing Mechanisms'. Insa Lyon, 5th Dec. 2024. DOI: 10.1145/3528535.3565240.URL10.1145/3528535.3565240.. URL: https://hal.science/tel-04874777 (cit. on p. 7).

[32] S. Pélissier. 'Privacy-preserving communications for the IoT'. INSA de Lyon, 27th Sept. 2024. URL: https://theses.hal.science/tel-04846788 (cit. on p. 7).

Reports & preprints

[33] M. Cunche. *Standardised Messenger Audit*. European Data Protection Board (EDPB), May 2024. URL: https://hal.science/hal-04840617 (cit. on pp. 18, 26).

[34] C. Santos, N. Bielova, V. Roca, M. Cunche, G. Mertens, K. Kubicek and H. Haddadi. *Feedback to the European Data Protection Board's Guidelines 2/2023 on Technical Scope of Art. 5(3) of ePrivacy Directive*. Inria & Université Cote d'Azur, CNRS, I3S, Sophia Antipolis, France, 18th Jan. 2024. DOI: 10.48550/arXiv.2402.02877. URL: https://inria.hal.science/hal-04437008 (cit. on p. 18).

Other scientific publications

[35] N. Bielova, L. Litvine, A. Nguyen, M. Chammat, V. Toubiana and E. Hary. *The Effect of Design Patterns on (Present and Future) Cookie Consent Decisions: Supplemental Materials*. Philadelphia, United States, 14th Aug. 2024. URL: https://inria.hal.science/hal-04235032.

[36] A. Lodie. *Case C-479/22 P, Case C-604/22 and the limitation of the relative approach of the definition of 'personal data' by the ECJ*. 25th Mar. 2024. URL: https://hal.science/hal-04609263 (cit. on p. 28).

[37] G. Mertens, N. Bielova, V. Roca and C. Santos. ‘Poster: Client-Side GTM Tags: hidden data leaks and potential violations of GDPR’. In: PETS 2024 - 24th Privacy Enhancing Technologies Symposium. Bristol, United Kingdom, 2024, pp. 1–1. URL: https://hal.science/hal-04699754.