

2025 Activity Report

RESEARCH CENTRES: Inria Centre at Université Grenoble Alpes
Inria Lyon Centre

IN PARTNERSHIP WITH: Institut national des sciences appliquées de Lyon

Project-Team

PRIVATICS

Privacy Models, Architectures and Tools for the
Information Society

In collaboration with Centre d'innovation en télécommunications et intégration
de services



Project-Team PRIVATICS

Creation of the Project-Team: 2014 July 01

Each year, Inria research teams publish an Activity Report presenting their work and results over the reporting period. These reports follow a common structure, with some optional sections depending on the specific team. They typically begin by outlining the overall objectives and research programme, including the main research themes, goals, and methodological approaches. They also describe the application domains targeted by the team, highlighting the scientific or societal contexts in which their work is situated. The reports then present the highlights of the year, covering major scientific achievements, software developments, or teaching contributions. When relevant, they include sections on software, platforms, and open data, detailing the tools developed and how they are shared. A substantial part is dedicated to new results, where scientific contributions are described in detail, often with subsections specifying participants and associated keywords. Finally, the Activity Report addresses funding, contracts, partnerships, and collaborations at various levels, from industrial agreements to international cooperations. It also covers dissemination and teaching activities, such as participation in scientific events, outreach, and supervision. The document concludes with a presentation of scientific production, including major publications and those produced during the year.

Keywords

Computer sciences and digital sciences

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

Contents

Project-Team PRIVATICS	1
1 Team members, visitors, external collaborators	6
2 Overall objectives	7
2.1 Context and overall objectives	7
3 Research program	8
4 Application domains	8
5 Social and environmental responsibility	9
5.1 Environmental impacts of research results	9
5.2 Societal impacts of research results	9
6 Highlights of the year	11
6.1 Awards	11
7 Latest software developments, platforms, open data	11
7.1 Latest software developments	11
7.1.1 gtm-eye	11
7.1.2 NLP Privacy	11
7.1.3 nlp-mem	12
7.1.4 ldp-audit	12
7.1.5 stetoscope	12
7.1.6 NoID-LLMs	12
8 New results	13
8.1 Research axis 1: AI	13
8.1.1 Group fairness under obfuscated sensitive information	13
8.1.2 SAAFL: Secure Aggregation for Label-Aware Federated Learning	13
8.1.3 Fair play for individuals, foul play for groups? Auditing anonymization’s impact on ML fairness	14
8.1.4 Inferring Communities of Interest in Collaborative Learning-based Recommender Systems	14
8.1.5 FADE: federated aggregation with discrimination elimination	14
8.1.6 PriviRec: Confidential and Decentralized Graph Filtering for Recommender Systems	15
8.1.7 Buffalo: A Practical Secure Aggregation Protocol for Buffered Asynchronous Federated Learning	15
8.1.8 Demo: Exploring Utility and Attackability Trade-offs in Local Differential Privacy	15
8.1.9 Bench-MIA: Towards automatic multi-lever benchmark construction for MIA evaluation	16
8.1.10 Exposing the Vulnerability of Decentralized Learning to Membership Inference Attacks Through the Lens of Graph Mixing	16
8.1.11 GRANITE: a Byzantine-Resilient Dynamic Gossip Learning Framework	17
8.1.12 Towards the Anonymization of the Language Modeling	17
8.2 Research axis 2: Web, smartphone, IoT, and wireless	17
8.2.1 Efficiently linking LoRaWAN identifiers through multi-domain fingerprinting	17
8.2.2 Towards Operational and Security Best Practices for DNS in the Internet of Things	18
8.2.3 You Can’t Trust Your Tag Neither: Privacy Leaks and Potential Legal Violations within the Google Tag Manager	18
8.2.4 k-scale: k-Anonymizing Millions of Trajectories	18
8.2.5 QRisk: Think Before You Scan QR codes	19
8.2.6 Retour d’expérience avec un jeu-débat pour sensibiliser à l’IA pour la surveillance sanitaire	19

8.3	Research axis 3: User empowerment	19
8.3.1	Leveraging interdisciplinary methods for evidence collection in enforcement: Dark patterns as a case study	19
8.3.2	Johnny Can't Revoke Consent Either: Measuring Compliance of Consent Revocation on the Web	20
8.3.3	Towards Key Contributing Factors in Identifying Dark Pattern Autonomy Violations under the EU Digital Services Act	20
8.3.4	"I'm Not for Sale" - Perceptions and Limited Awareness of Privacy Risks by Digital Natives About Location Data	21
8.4	Research axis 4: Legal	21
8.4.1	Usable and Lawful: Can Consent Be Both?	21
8.4.2	Understanding the scope of Article 25 of the DSA in regulating dark patterns	22
8.4.3	Feedback to the EU Commission's Call for evidence for an impact assessment — Ares(2025)5829481 on the Digital Fairness Act	22
8.4.4	Feedback to the European Data Protection Board's Guidelines 3/2025 on the interplay between the DSA and the GDPR (Version 1.1) - Advertisement	22
8.4.5	The Impact of EU Digital Regulation on the Protection of Vulnerability Disclosure Researchers	22
8.4.6	PIA : Enseigner la protection des données personnelles dans l'interdisciplinarité	23
8.5	Research axis 5: Applied Cryptography	23
8.5.1	Online/Offline Digital Signatures: A Systematic Literature Review	23
9	Bilateral contracts and grants with industry	23
9.1	Bilateral contracts with industry	23
10	Partnerships and cooperations	24
10.1	International initiatives	24
10.1.1	Inria associate team not involved in an IIL or an international program	24
10.2	International research visitors	24
10.2.1	Visits of international scientists	24
10.3	European initiatives	25
10.3.1	Digital Europe	25
10.4	National initiatives	25
10.4.1	ANR	25
10.4.2	Inria Exploratory Action (AEx)	28
10.4.3	Others	29
10.5	Public policy support	29
11	Dissemination	29
11.1	Promoting scientific activities	30
11.1.1	Scientific events: organisation	30
11.1.2	Scientific events: selection	30
11.1.3	Journal	31
11.1.4	Invited talks	31
11.1.5	Leadership within the scientific community	31
11.1.6	Scientific expertise	32
11.1.7	Research administration	32
11.2	Teaching - Supervision - Juries - Educational and pedagogical outreach	32
11.2.1	Juries	32
11.2.2	Educational and pedagogical outreach	32
11.3	Popularization	33
11.3.1	Others science outreach relevant activities	33

12 Scientific production	33
12.1 Major publications	33
12.2 Publications of the year	33

1 Team members, visitors, external collaborators

Research Scientists

- Vincent Roca [Team leader, INRIA, Researcher, HDR]
- Nataliia Bielova [INRIA, Senior Researcher, until Nov 2025, HDR]
- Claude Castelluccia [INRIA, Senior Researcher, HDR]
- Heber Hwang Arcolezi [INRIA, ISFP]
- Cedric Lauradoux [INRIA, Researcher]
- Mohamed Maouche [INRIA, ISFP]
- Cristiana Teixeira Santos [INRIA, Starting Research Position, from Jun 2025 until Jul 2025]

Faculty Members

- Antoine Boutet [INSA LYON, Associate Professor, HDR]
- Mathieu Cunche [INSA LYON, Professor, HDR]
- Clementine Gritti [INSA LYON, Chair, Professeur Junior]

Post-Doctoral Fellows

- Sanju Ahuja [INRIA, Post-Doctoral Fellow, until Oct 2025]
- Karel Kubicek [INRIA, Post-Doctoral Fellow, until Mar 2025]
- Paul Lachat [INRIA, Post-Doctoral Fellow, from Sep 2025]
- Abhishek Mishra [INRIA, Post-Doctoral Fellow]

PhD Students

- Ivan Baheux-Blin [MURENA SAS, CIFRE]
- Imen Bajar [INRIA, from Nov 2025]
- Teodora Curelariu [UGA]
- Marwa El Kamil [INRIA, from Dec 2025]
- Pascal Engelibert [INSA LYON, from Oct 2025]
- Jules Marmier [INRIA, until May 2025]
- Gilles Mertens [INRIA]
- Alix Ntoutoume Nzame [OPEN SEZAM SAS, CIFRE]
- Zhan Xu [INRIA, from Nov 2025]

Technical Staff

- Mohamed Bechorfa [INRIA, Engineer, until Jan 2025]
- Lucas Magnana [INRIA, Engineer]

Interns and Apprentices

- Taha Aftiss [INRIA, Intern, from May 2025 until Jul 2025]
- Vinicius Gabriel Angelozzi Verona De Resende [INRIA, Intern, from May 2025 until Sep 2025]
- Vinicius Gabriel Angelozzi Verona De Resende [INRIA, Intern, until Apr 2025]
- Tao Beaufiles [INRIA, Intern, from Feb 2025 until Jul 2025]
- Mohammed Amine Benrahmoune [INRIA, Intern, from Jun 2025 until Aug 2025]
- Soraya Djerrab [INSA LYON, Intern, from Feb 2025 until Aug 2025]
- Youssef El Atia [INRIA, Intern, from May 2025 until Jul 2025]
- Duarte Miguel Ferreira Moreira Da Silva [INRIA, Intern, from May 2025 until Jul 2025]
- Eseosa Omorogieva [INRIA, Intern, from May 2025 until Jun 2025]
- Eseosa Omorogieva [INRIA, Intern, until Apr 2025]
- Annika Sauer [INRIA, Intern, from Feb 2025 until Jun 2025]
- Alice Vettier [INRIA, Intern, from May 2025 until Jun 2025]
- Alice Vettier [INRIA, Intern, from Mar 2025 until Apr 2025]
- Esteban Wybouw [INRIA, Intern, from Jun 2025 until Aug 2025]

Administrative Assistants

- Marie-Anne Dauphin-Rizzi [INRIA]
- Helen Pouchot-Rouge-Blanc [INRIA]

Visiting Scientists

- Salvatore Della Torca [Univ Bergamo, from Apr 2025 until Jul 2025]
- Cristiana Santos [Utrecht University, from Jun 2025 until Jul 2025, Assistant Professor]

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."

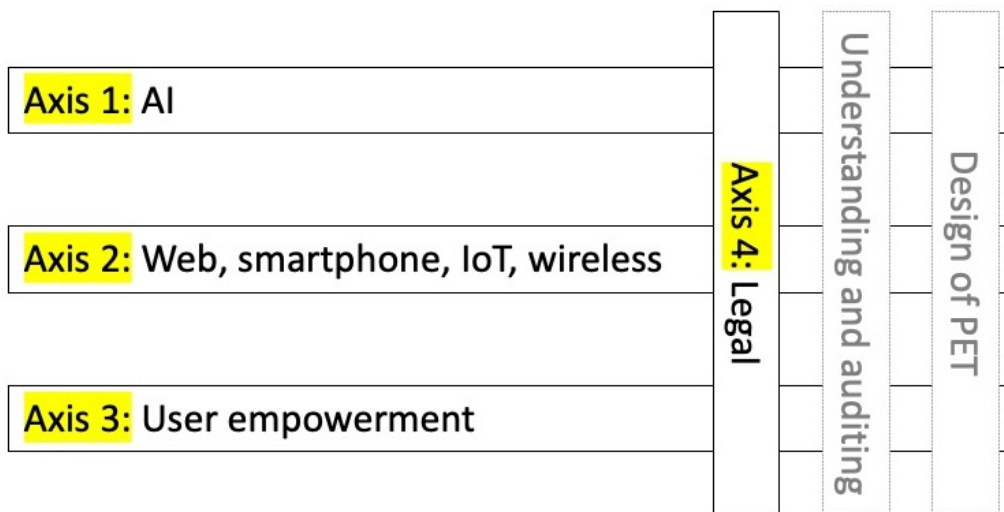
Making the world – a little bit – better: Privacy is thus essential to protect individuals, for instance against potential misuses of personal data. Privacy is also essential to protect the society, as has been highlighted by the misuse of personal data in order to surreptitiously influence voters in 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 axis 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 implementation 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 GAFAM 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 (comité d'évaluation des risques légaux et éthiques);
- Cédric Lauradoux represents the Inria COERLE (comité 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.
- Antoine Boutet is member of the *Commission Intelligence Artificielle (CIA)* of the HCL.

6 Highlights of the year

6.1 Awards

Nataliia Bielova has received two awards in 2025:

- [French Academy of Science Lovelace-Babbage Award](#), October 2025.
- [CNIL Inria Privacy Award](#), by the French Data Protection Authority (CNIL) and Inria, July 2025.

Héber Hwang Arcolezi has received a “Notable Reviewer Award” at USENIX Security 2025.

7 Latest software developments, platforms, open data

7.1 Latest software developments

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: This work is associated to a 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>

Publication: hal-04299405

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

7.1.5 stetoscope

Name: stetoscope

Keywords: Web, Mobile application, Personalized systems

Functional Description: Stetoscope is a platform which aims to analyze content personalization and to audit Web and mobile platforms. This application adopts a participatory approach by involving users (i.e., capture screenshots from mobile devices) to better understand platform practices (e.g., price or search personalization, price participation, incentive mechanisms, information manipulation) and raise their awareness of the issues. An administration dashboard also allows for the management of data collection campaigns. Due to the nature of the connected data (visual information in the form of screenshots), this tool can be adapted to multiple use cases.

Contact: Antoine Boutet

7.1.6 NoID-LLMs

Name: NoID-LLMs

Keywords: Privacy, LLM

Functional Description: Off-the-shelf LLMs are often specialized for specific datasets (e.g., medical reports). The memorization of sensitive information by the LLM is a source of risk that can lead to subsequent privacy leakage, whether the model is used in a black box or shared. The NoID-LLMs library proposes a privacy-preserving specialization scheme that prevents the memorization of identifiers (both direct and indirect). This library also provides tools for sanitizing LLMs. Specifically, this library leverages unlearning methods to forget sensitive information that might have been memorized by a model specialized without specific safeguards.

Contact: Antoine Boutet

8 New results

8.1 Research axis 1: AI

8.1.1 Group fairness under obfuscated sensitive information

Participants: Héber Arcolezi, et al..

In the era of Big Data, the development of artificial intelligence (AI) systems presents both opportunities and challenges, particularly concerning privacy and fairness. While differential privacy (DP) has emerged as a robust methodology for preserving privacy in real-world applications, its local variant (LDP) specifically addresses trust issues by removing the reliance on a centralized server. Equally critical, conducting fairness audits of AI systems helps identify and mitigate discriminatory outcomes in machine learning. Although the relationship between DP and fairness is inherently multifaceted, this paper offers a detailed empirical examination of how collecting multi-dimensional sensitive attributes under LDP affects fairness in binary classification tasks. Our findings reveal that LDP can slightly improve fairness without substantially degrading model performance—challenging the notion that DP necessarily exacerbates unfairness. We demonstrate these results by evaluating seven state-of-the-art LDP protocols on three benchmark datasets, using established group fairness metrics. Moreover, we propose a novel privacy budget allocation scheme that incorporates varying domain sizes of sensitive attributes, achieving a superior privacy–utility–fairness trade-off compared to existing solutions.

Related publication: [6]

8.1.2 SAAFL: Secure Aggregation for Label-Aware Federated Learning

Participants: Clémentine Gritti, et al..

Secure aggregation (SA) has emerged as a vital component of federated learning (FL), enabling collaborative training of a global machine learning model while safeguarding the privacy of clients' local datasets. Most existing SA protocols implement the privacy preserving variant of federated averaging (FedAvg) as the aggregation technique and assume independent and identically distributed (IID) datasets across clients. This assumption makes FedAvg unsuitable for non-IID scenarios, where variations in client datasets lead to less effective global model. We propose SAAFL, a SA protocol specifically designed for non-IID settings and more specifically for the recently proposed federated label-aware aggregation (FedLA) protocol. SAAFL computes the weighted average of clients' inputs where weights depend on the label distributions and should remain confidential. SAAFL is resilient to client dropouts and supports client selection. Our experimental results show that it achieves comparable model accuracy with FedLA and remains efficient in terms of computation and communication.

Related publication: [13]

8.1.3 Fair play for individuals, foul play for groups? Auditing anonymization’s impact on ML fairness

Participants: Héber Arcolezi, et al..

Machine learning (ML) algorithms are heavily based on the availability of training data, which, depending on the domain, often includes sensitive information about data providers. This raises critical privacy concerns. Anonymization techniques have emerged as a practical solution to address these issues by generalizing features or suppressing data to make it more difficult to accurately identify individuals. Although recent studies have shown that privacy-enhancing technologies can influence ML predictions across different subgroups, thus affecting fair decision-making, the specific effects of anonymization techniques, such as k-anonymity, l-diversity, and t-closeness, on ML fairness remain largely unexplored. In this work, we systematically audit the impact of anonymization techniques on ML fairness, evaluating both individual and group fairness. Our quantitative study reveals that anonymization can degrade group fairness metrics by up to fourfold. Conversely, similarity-based individual fairness metrics tend to improve under stronger anonymization, largely as a result of increased input homogeneity. By analyzing varying levels of anonymization across diverse privacy settings and data distributions, this study provides critical insights into the trade-offs between privacy, fairness, and utility, offering actionable guidelines for responsible AI development. Our code is publicly available.

Related publication: [14]

8.1.4 Inferring Communities of Interest in Collaborative Learning-based Recommender Systems

Participants: Mohamed Maouche, et al..

Collaborative-learning-based recommender systems, such as those employing Federated Learning (FL) and Gossip Learning (GL), allow users to train models while keeping their history of liked items on their devices. While those methods were seen as promising for enhancing privacy, recent research has shown that collaborative learning can be vulnerable to various privacy attacks. In this paper, we propose a novel attack called Community Inference Attack (CIA), which enables an adversary to identify community members based on a set of target items. What sets CIA apart is its efficiency: it operates at a low computational cost by eliminating the need for training surrogate models. Instead, it uses a comparison-based approach, inferring sensitive information by comparing users’ models rather than targeting any specific individual model. To evaluate the effectiveness of CIA, we conduct experiments on three real-world recommendation datasets using two recommendation models under both federated and gossip-like settings. The results demonstrate that CIA can be up to 10 times more accurate than random guessing. Additionally, we evaluate two mitigation strategies: Differentially Private Stochastic Gradient Descent (DP-SGD) and a Share less policy, which involves sharing fewer, less sensitive model parameters. Our findings suggest that the Share less strategy offers a better privacy-utility trade-off, especially in GL.

Related publication: [15]

8.1.5 FADE: federated aggregation with discrimination elimination

Participants: Héber Arcolezi, et al..

In this work, we investigate how unfair updates with opposing biases can cancel each other out during aggregation in federated learning (FL), leading to a fairer overall model from a group fairness perspective. We analytically and empirically analyze this Federated Aggregation with Discrimination Elimination (FADE) phenomenon, considering both linear and nonlinear models. In addition, we build on this observation and introduce two novel fairness-aware FL aggregation strategies. The first strategy, FADE-OptW, uses sequential optimization to optimize weights assigned to each client based on their fairness levels. The

second approach, FADE-SSP, identifies the optimal subset of clients that minimizes the weighted average fairness level at each round along the convergence path, and for a given metric. Our experiments demonstrate significant improvements in fairness, achieving up to a 60% reduction in discrimination compared to standard FedAvg-based FL. We achieve these gains while maintaining the model’s predictive performance on highly heterogeneous client data distributions.

Related publication: [16]

8.1.6 PriviRec: Confidential and Decentralized Graph Filtering for Recommender Systems

Participants: Mohamed Maouche, et al..

Recent advances in recommender systems have shown that relying on graph filters, such as the normalized item-item adjacency matrix and the ideal low-pass filter yields competitive performance and scales better than Graph Convolutional Networks-based solutions. However, these solutions require centralizing user data, which raises concerns over data privacy, security, and the monopolization of user data by a few actors. To address those concerns, we propose PriviRec and PriviRec-k, two complementary recommendation frameworks. In PriviRec, we show that it is possible to decompose widely used filters so that they can be computed in a distributed setting using Secure Aggregation and a distributed version of the Randomized Power Method, without revealing individual users contributions. PriviRec-k extends this approach by having users securely aggregate low-rank projections of their contributions, enabling a tunable balance between communication overhead and recommendation accuracy. We demonstrate theoretically as well as experimentally on Gowalla, Yelp2018, and Amazon-Book that our methods achieve performance comparable to centralized state-of-the-art recommender systems and superior to decentralized ones, while preserving confidentiality and low communication and computational overheads.

Related publication: [24]

8.1.7 Buffalo: A Practical Secure Aggregation Protocol for Buffered Asynchronous Federated Learning

Participants: Clémentine Gritti, et al..

Federated Learning (FL) has become a crucial framework for collaboratively training Machine Learning (ML) models while ensuring data privacy. Traditional synchronous FL approaches, however, suffer from delays caused by slower clients (called stragglers), which hinder the overall training process. Specifically, in a synchronous setting, model aggregation happens once all the intended clients have submitted their local updates to the server. To address these inefficiencies, Buffered Asynchronous FL (BAsyncFL) was introduced, allowing clients to update the global model as soon as they complete local training. In such a setting, the new global model is obtained once the buffer is full, thus removing synchronization bottlenecks. Despite these advantages, existing Secure Aggregation (SA) techniques-designed to protect client updates from inference attacks-rely on synchronized rounds, making them unsuitable for asynchronous settings. In this paper, we present Buffalo, the first practical SA protocol tailored for BAsyncFL. Buffalo leverages lattice-based encryption to handle scalability challenges in large ML models and introduces a new role, the assistant, to support the server in securely aggregating client updates. To protect against an actively corrupted server, we enable clients to verify that their local updates have been correctly integrated into the global model. Our comprehensive evaluation-incorporating theoretical analysis and real-world experiments on benchmark datasets-demonstrates that Buffalo is an efficient and scalable privacy-preserving solution in BAsyncFL environments.

Related publication: [25]

8.1.8 Demo: Exploring Utility and Attackability Trade-offs in Local Differential Privacy

Participants: Abhishek Mishra, Héber Arcolezi, et al..

Local Differential Privacy (LDP) provides strong, formal privacy guarantees without requiring a trusted curator, making it a promising approach for privacy-preserving data collection and analysis. However, despite extensive research, practitioners may struggle to understand how to tune LDP parameters and anticipate the impact on data utility and attack risks for their specific scenarios. To address this gap, we demonstrate LDP-Toolbox, the first interactive, web-based toolbox (implemented in Python) that enables practical, analytical visualization of trade-offs between privacy loss (ϵ), utility loss, and vulnerability to attacks. The toolbox supports exploration of these trade-offs using real-world datasets from different domains; in this demonstration, we focus on discrete personal attributes and location-based scenarios. By providing intuitive, visual insights, LDP-Toolbox lowers the barrier to deploying LDP in real applications and helps bridge the gap between theoretical guarantees and practical adoption. The toolbox is open-source on [PyPI](#) and a video is available on our [GitHub repository](#).

Related publication: [26]

8.1.9 Bench-MIA: Towards automatic multi-lever benchmark construction for MIA evaluation

Participants: Héber Arcolezi, et al..

The lack of transparency in large language models (LLMs) training data has created concerns regarding unauthorized data use. Membership Inference Attacks (MIA) offer a method for detecting potential copyright violations and privacy breaches by determining whether specific data points were included during training. However, existing MIA methods face reliability and generalization challenges due to biases in validation datasets. Since various LLMs are trained on different, often proprietary datasets, creating a single universal dataset and evaluation framework is not feasible. In this work-in-progress paper, we introduce bench-MIA, an approach that automatically constructs tailored benchmarks for MIA evaluation. Bench-MIA leverages the concept of lever to systematically generate evaluation datasets addressing key sources of bias. By providing appropriate evaluation baselines for each dataset, bench-MIA facilitates more accurate and fair MIA assessments. Preliminary experimental results demonstrate the impact of these levers on MIA evaluation and motivate further validation.

Related publication: [28]

8.1.10 Exposing the Vulnerability of Decentralized Learning to Membership Inference Attacks Through the Lens of Graph Mixing

Participants: Mohamed Maouche, et al..

The primary promise of decentralized learning is to allow users to engage in the training of machine learning models in a collaborative manner while keeping their data on their premises and without relying on any central entity. However, this paradigm necessitates the exchange of model parameters or gradients between peers. Such exchanges can be exploited to infer sensitive information about training data, which is achieved through privacy attacks (e.g., Membership Inference Attacks – MIA). In order to devise effective defense mechanisms, it is important to understand the factors that increase/reduce the vulnerability of a given decentralized learning architecture to MIA. In this study, we extensively explore the vulnerability to MIA of various decentralized learning architectures by varying the graph structure (e.g., number of neighbors), the graph dynamics, and the aggregation strategy, across diverse datasets and data distributions. Our key finding, which to the best of our knowledge we are the first to report, is that the vulnerability to MIA is heavily correlated to (i) the local model mixing strategy performed by each node upon reception of models from neighboring nodes and (ii) the global mixing properties of the communication graph. We illustrate these

results experimentally using four datasets and by theoretically analyzing the mixing properties of various decentralized architectures. We also empirically show that enhancing mixing properties is highly beneficial when combined with other privacy-preserving techniques such as Differential Privacy. Our paper draws a set of lessons learned for devising decentralized learning systems that reduce by design the vulnerability to MIA.

Related publication: [29]

8.1.11 GRANITE: a Byzantine-Resilient Dynamic Gossip Learning Framework

Participants: Mohamed Maouche, et al..

Gossip Learning (GL) is a decentralized learning paradigm where users iteratively exchange and aggregate models with a small set of neighboring peers. Recent GL approaches rely on dynamic communication graphs built and maintained using Random Peer Sampling (RPS) protocols. Thanks to graph dynamics, GL can achieve fast convergence even over extremely sparse topologies. However, the robustness of GL over dynamic graphs to Byzantine (model poisoning) attacks remains unaddressed especially when Byzantine nodes attack the RPS protocol to scale up model poisoning. We address this issue by introducing GRANITE, a framework for robust learning over sparse, dynamic graphs in the presence of a fraction of Byzantine nodes. GRANITE relies on two key components (i) a History-aware Byzantine-resilient Peer Sampling protocol (HaPS), which tracks previously encountered identifiers to reduce adversarial influence over time, and (ii) an Adaptive Probabilistic Threshold (APT), which leverages an estimate of Byzantine presence to set aggregation thresholds with formal guarantees. Empirical results confirm that GRANITE maintains convergence with up to 30% Byzantine nodes, improves learning speed via adaptive filtering of poisoned models and obtains these results in up to 9 times sparser graphs than dictated by current theory.

Related publication: [32]

8.1.12 Towards the Anonymization of the Language Modeling

Participants: Antoine Boutet, et al..

Rapid advances in Natural Language Processing (NLP) have revolutionized many fields, including healthcare. However, these advances raise significant privacy concerns, especially when pre-trained models fine-tuned and specialized on sensitive data can memorize and then expose and regurgitate personal information. This work presents a privacy-preserving language modeling approach to address the problem of language models anonymization, and thus promote their sharing. Specifically, we propose both a Masking Language Modeling (MLM) methodology to specialize a BERT-like language model, and a Causal Language Modeling (CLM) methodology to specialize a GPT-like model that avoids the model from memorizing direct and indirect identifying information present in the training data. We have comprehensively evaluated our approaches using a medical dataset and compared them against different baselines. Our results indicate that by avoiding memorizing both direct and indirect identifiers during model specialization, our masking and causal language modeling schemes offer a good tradeoff for maintaining high privacy while retaining high utility.

Related publication: [36]

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

8.2.1 Efficiently linking LoRaWAN identifiers through multi-domain fingerprinting

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

LoRaWAN is a leading IoT technology worldwide, increasingly integrated into pervasive computing environments through a growing number of sensors in various industrial and consumer applications. Although its security vulnerabilities have been extensively explored in the recent literature, its ties to human activities warrant further privacy research. Existing device identification and activity inference attacks are only effective with a stable identifier. We find that the identifiers in LoRaWAN exhibit high variability, and more than half of the devices use them for less than a week. For the first time in the literature, we explore the feasibility of device fingerprinting in LoRaWAN, allowing long-term device linkage, i.e. associating various identifiers of the same device. We introduce a novel holistic fingerprint representation utilizing multiple domains, namely content, timing, and radio information, and present a machine learning-based solution for linking identifiers. Through a large-scale experimental evaluation based on realworld datasets containing up to 41 million messages, we study multiple scenarios, including an attacker with limited resources. We reach 0.98 linkage accuracy, underscoring the need for privacy-preserving measures. We showcase countermeasures including payload padding, random delays, and radio signal modulation, and conclude by assessing their impact on our fingerprinting solution.

Related publication: [9]

8.2.2 Towards Operational and Security Best Practices for DNS in the Internet of Things

Participants: Abhishek Mishra, Mathieu Cunche, et al..

The Domain Name System (DNS) is vital for Internet operation, but its lack of standards for Internet of Things (IoT) devices raises security and reliability concerns. This paper investigates inconsistencies in IoT DNS operations, revealing both security risks and irregular behaviors. We analyze DNS on a large IoT testbed through passive traffic inspection and active testing, uncovering serious anomalies. Our findings highlight vulnerabilities to cache poisoning, fingerprinting, and DoS attacks. We assess standardization gaps in IoT DNS security and move towards proposing guidelines to enhance resilience.

Related publication: [19]

8.2.3 You Can't Trust Your Tag Neither: Privacy Leaks and Potential Legal Violations within the Google Tag Manager

Participants: Gilles Mertens, Nataliia Bielova, Vincent Roca, Cristiana Santos.

Tag Management Systems (TMS) were developed in order to support website Publishers in installing multiple third-party JavaScript scripts (Tags) on their websites. Google has proposed its own TMS called "Google Tag Manager" (GTM) that is currently present on 52% of the top 1 million most popular websites. However, GTM has not yet been thoroughly evaluated by the academic research community. In this work, we study, for the first time, the Tags provided within the GTM system. Our methodology consists in installing Tags in isolation to analyze the types of data that Tags collect and contrast them to the legal and technical documentation, in collaboration with a legal expert. Across three studies – in-depth analysis of 6 Tags, automated analysis of 718 Tags, and analysis of Google "Consent Mode" – we discover multiple hidden data leaks, incomplete and diverging declarations, undisclosed third-parties and cookies, personal data sharing without consent and we further identify potential legal violations within EU Data Protection law.

Related publication: [20]

8.2.4 k-scale: k-Anonymizing Millions of Trajectories

Participants: Abhishek Mishra, et al..

Trajectory datasets collected by network operators and service providers offer detailed information about individual mobility and have wide application in business and research. However, managing such data raises privacy risks, as the unique movement patterns of individuals pose significant re-identification risks and make common countermeasures like pseudonymization ineffective. The privacy-preserving data publishing (PPDP) of trajectory datasets that maintains post-anonymization accuracy and truthfulness is an open problem -especially for large datasets with millions of records like those gathered by major actors in the telco ecosystem. We close this gap with *k-scale*, a framework that implements *k*-anonymity in massive mobile user trajectory datasets, removing uniqueness while safeguarding accuracy at the record level. Not only *k-scale* is the first model capable of scaling *k*-anonymization to a dataset of one million trajectories, but it does so while also outperforming state-of-the-art methods for trajectory data publishing in terms of preserved data quality, which we prove in real-world massive datasets and applications.

Related publication: [22]

8.2.5 QRisk: Think Before You Scan QR codes

Participants: Abhishek Mishra, Mathieu Cunche, et al..

QR codes are pervasive in modern digital interactions, but despite their convenience, they pose significant privacy risks that are often underestimated. For instance, privacy issues escalate when scanned URLs trigger HTTP redirections involving QR URL shorteners and third-party domains, exposing user data to external entities. However, a comprehensive study of the privacy implications of QR code interactions concerning cookie exploitation and query strings remains lacking in the literature. To address this, we collected a dataset of 860 QR codes over a two-year period from France, China, Austria, India, and Canada, to analyze the privacy risks associated with QR code usage. In this paper, QRisk, we find that 39.2% of redirected URLs set cookies, including tracking, analytics, and advertising cookies, enabling potential cross-session behavioral profiling. Additionally, over 25% of QR URLs embed query strings that not only contain sensitive user identifiers but also carry information such as location data, leading to user profiling and social link inference.

Related publication: [23]

8.2.6 Retour d'expérience avec un jeu-débat pour sensibiliser à l'IA pour la surveillance sanitaire

Participants: Cédric Lauradoux, et al..

L'importance de l'Intelligence Artificielle dans la société, et des décisions qui lui sont déléguées (accès aux formations supérieures, détermination de peines de prison, conduite autonome de véhicules, etc) nécessite d'éduquer la population à ses enjeux. Tout le monde ne peut pas avoir une connaissance technique précise sur l'IA, mais comme souligné par l'UNESCO, il est essentiel que la population ait une compréhension basique du fonctionnement de ces algorithmes pour choisir en connaissance de cause de les utiliser ou pas. Pour cela nous avons développé un jeu sérieux à destination des lycées, sous la forme d'un débat citoyen ayant pour but de choisir une solution d'IA pour contrôler une épidémie. Cet article présente un retour d'expérience de l'utilisation de ce jeu en classe. Nous décrivons le fonctionnement d'une séance, les résultats qualitatifs et quantitatifs des premières sessions animées, ainsi que les biais et limites de cette activité.

Related publication: [39]

8.3 Research axis 3: User empowerment

8.3.1 Leveraging interdisciplinary methods for evidence collection in enforcement: Dark patterns as a case study

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

"Dark patterns" are manipulative, deceptive design practices deployed in online services to influence users' decisions towards undesired or negative outcomes. Interdisciplinary by nature, dark patterns implicate concepts of autonomy and choice from law, human behaviour from the psychology and social science disciplines, and design and human-computer interaction (HCI) from technical fields and industry. A body of enforcement actions and regulatory fines worldwide as discussed within this article comprise a growing effort to minimise the impact of dark patterns. However, despite this regulatory momentum, it remains unknown to what extent scientific research methods and evidence types may influence regulatory decisions, which is relevant for effective evidence-based enforcement. As such, dark patterns present a case study for reflecting upon narrowing the academic-enforcement divide. Our team spans design, HCI, computer science, and law, and examines investigatory methodologies towards insight for strengthening collaboration between scholars and regulators. This interdisciplinary work considers investigatory methods from both academia and industry, then as inferred from dark patterns enforcement cases to relate methods used by both groups. We discuss challenges and opportunities for tightening the gap between researchers and regulators, and propose suggestions for both scholars and enforcers to tighten feedback loops. We additionally highlight informal investigation methods as an opportunity to strengthen collaboration.

Related publication: [7]

8.3.2 Johnny Can't Revoke Consent Either: Measuring Compliance of Consent Revocation on the Web

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

The EU General Data Protection Regulation (GDPR) requires websites to facilitate the right to revoke consent from Web users. Prior works have examined consent management by auditing that user choices are correctly stored, and comparing cookies set upon acceptance versus rejection to assess compliance. While these studies measured compliance of consent with respect to the various consent requirements, no prior work has studied consent revocation on the Web. Therefore, it is unclear how difficult it is to revoke consent on the websites' interfaces, and whether the revoked consent is properly stored and communicated behind the user interface. Our work aims to fill this gap by measuring compliance of consent revocation on the Web on Tranco's top-200 websites. We found that 19.87% of websites make it difficult for users to revoke consent throughout different interfaces, 20.5% of websites require more effort than acceptance, and 2.48% do not provide consent revocation at all, thus violating EU legal requirements for valid consent. 57.5% websites do not delete the cookies after consent revocation enabling continuous illegal processing of users' data. Further, we analyzed 281 websites implementing the IAB Europe Transparency and Consent Framework, and found 22 websites that store a positive consent despite user's revocation. Surprisingly, we found that on 101 websites, third parties that have received consent upon user's acceptance, are not informed of revocation, leading to the illegal processing of users' data by such third parties according to EU laws. Our findings emphasize the need for improved legal compliance of consent revocation, and proper, consistent, and uniform implementation of revocation communication to third-parties.

Related publication: [8]

8.3.3 Towards Key Contributing Factors in Identifying Dark Pattern Autonomy Violations under the EU Digital Services Act

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

Dark patterns refer to design practices which undermine users' ability to make autonomous and informed choices in relation to digital systems. The recent EU Digital Services Act (DSA) aims to protect users from such dark patterns and their effects. DSA Article 25 prohibits three autonomy violation types: deception, manipulation and distortion/impairment. However, for regulation of dark patterns, it is important to reason about why an observed design practice constitutes a particular autonomy violation type, to show that it indeed

violates the DSA. In this work-in-progress, two experts (with HCI, CS and legal background) mapped 59 known dark patterns onto these three autonomy violation types. We then analysed our rationale for this mapping to identify eight design factors which can help determine the dark pattern autonomy violation(s). Our analysis aims to situate existing dark patterns knowledge within the DSA legal framework, to support regulation and compliance of such design practices.

Related publication: [12]

8.3.4 "I'm Not for Sale" - Perceptions and Limited Awareness of Privacy Risks by Digital Natives About Location Data

Participants: Antoine Boutet, et al..

Although mobile devices benefit users in their daily lives in numerous ways, they also raise several privacy concerns. For instance, they can reveal sensitive information that can be inferred from location data. This location data is shared through service providers as well as mobile applications. Understanding how and with whom users share their location data as well as users' perception of the underlying privacy risks, are important notions to grasp in order to design usable privacy-enhancing technologies. In this work, we perform a quantitative and qualitative analysis of smartphone users' awareness, perception and self-reported behavior towards location data-sharing through a survey of n=99 young adult participants (i.e., digital natives). We compare stated practices with actual behaviors to better understand their mental models, and survey participants' understanding of privacy risks before and after the inspection of location traces and the information that can be inferred therefrom. Our empirical results show that participants have risky privacy practices: about 54% of participants underestimate the number of mobile applications to which they have granted access to their data, and 33% forget or do not think of revoking access to their data. Furthermore, most of the participants do not have a realistic perception of privacy risks and have generally heard little about privacy-related scandals. Also, by using a demonstrator to perform inferences from location data, we observe that slightly more than half of participants (57%) are surprised by the extent of potentially inferred information, and that 47% intend to reduce access to their data via permissions as a result of using the demonstrator. Last, a majority of participants have little knowledge of the tools to better protect themselves, but are nonetheless willing to follow suggestions to improve privacy (51%). Educating people, including digital natives, about privacy risks through transparency tools seems a promising approach.

Related publication: [17]

8.4 Research axis 4: Legal

8.4.1 Usable and Lawful: Can Consent Be Both?

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

Under the GDPR, a valid consent must satisfy a number of requirements to comply with the General Data Protection Regulation (GDPR) and the ePrivacy Directive (ePD). The article evaluates the design of consent banners using a common and popular usability inspection method in human-computer interaction scholarship known as heuristic evaluation, which enables the researcher to identify challenges in technical provision and new generative opportunities for technical systems to better respond to legal requirements. Opportunities, challenges and tensions for a lawful and usable consent are identified through a novel application of usability heuristics to target the intersection of legal requirements and usability in the context of consent banners. These interpretations of the intersection of law and design may aid legal scholars in evaluating the lawfulness and usability of consent design strategies, while acknowledging the tensions and challenges among design and legal perspectives.

Related publication: [11]

8.4.2 Understanding the scope of Article 25 of the DSA in regulating dark patterns

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

The Digital Services Act (DSA) became directly applicable across the EU, explicitly codifying and prohibiting dark patterns in online interfaces for the first time in its Article 25. However, the DSA regulators across EU now are faced with an important challenge: how to reason about the prohibitions set out in the Article 25(1) and its Recital 67, since these provisions protect user autonomy, which is not a legal concept? And how to ensure that a given practice indeed deceives, manipulates or materially distorts or impairs the ability of the recipients to make free and informed decisions? Additionally, it is unclear how the interplay between the DSA Article 25(2) and the UCPD can be applied in practice. In this book chapter, we give answers to these questions by analysing the literature in philosophy and Human-Computer Interaction (HCI) domains. This body of knowledge has already provided several definitions to the concept of user autonomy. Drawing on it, we propose three autonomy violation types set out in the DSA Article 25(1) that should help regulators reason about the violations of Article 25(1) in practice. We also advocate for the usage of the most comprehensive ontology of dark patterns based on 5 academic and 5 regulatory taxonomies, and map dark patterns in the DSA Article 25(3) to the ontology, demonstrating the usefulness of the ontology when reasoning about dark patterns. Finally, we contribute to the discussion on the interplay between the DSA Article 25 and UCPD and propose three perspectives that regulators can take to resolve ambiguities between the two laws.

Related report: [santos:hal-05308131](#)

8.4.3 Feedback to the EU Commission’s Call for evidence for an impact assessment — Ares(2025)5829481 on the Digital Fairness Act

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

In response to the EU Commission’s Call for Evidence, we present our inputs organized by topics, drawing from academic knowledge from the fields of Law, Design, and HumanComputer Interaction within Computer Science. Our contributions are aligned with the Commission’s objectives and policy options presented below.

Related publication: [\[33\]](#)

8.4.4 Feedback to the European Data Protection Board’s Guidelines 3/2025 on the interplay between the DSA and the GDPR (Version 1.1) - Advertisement

Participants: Nataliia Bielova, et al..

Hereunder is our feedback to the EDPB Guidelines 3/2025 on the interplay between the DSA and the GDPR. Our comments are presented after a quotation from the proposed text in a box. The highlights in bold were added by the authors.

Related publication: [\[37\]](#)

8.4.5 The Impact of EU Digital Regulation on the Protection of Vulnerability Disclosure Researchers

Participants: Teodora Curelariu, et al..

While the European Union has taken important steps toward the institutionalisation of coordinated vulnerability disclosure, the protection of those who enable this process remains inadequate. The current

regulatory environment is marked by fragmentation, ambiguity, and uneven enforcement. Researchers face disproportionate risks, and this deters valuable contributions to collective cybersecurity. The experiences of the Netherlands and Belgium demonstrate that coherent legal frameworks, built on trust and mutual recognition, can foster a productive disclosure ecosystem without compromising security or legal certainty. As NIS 2 and the Cyber Resilience Act are implemented, the EU must embed strong safeguards to ensure a secure, ethical, and effective CVD system for Europe's digital future.

Related publication: [18]

8.4.6 PIA : Enseigner la protection des données personnelles dans l'interdisciplinarité

Participants: Antoine Boutet, et al..

Le module d'enseignement PIA (Privacy Impact Assessment), conjuguant informatique et droit, entend créer un dialogue concret entre étudiants en informatique et étudiants en droit afin d'envisager de manière novatrice la protection des données personnelles. Associant l'INSA de Lyon, la Faculté de droit de Nantes Université ainsi que la CNIL, le module PIA, dont la première séance a été organisée en février 2025, doit permettre aux étudiants d'acquérir des connaissances techniques et juridiques centrales pour leur cursus tout en les familiarisant avec l'interdisciplinarité et la nécessité d'engager un dialogue constructif au-delà de leur domaine de compétence.

Related publication: [27]

8.5 Research axis 5: Applied Cryptography

8.5.1 Online/Offline Digital Signatures: A Systematic Literature Review

Participants: Clémentine Gritti, et al..

In the rapidly evolving digital landscape, efficient cryptographic systems are increasingly critical. Digital Signature (DS) schemes are essential for ensuring data integrity and authenticity, particularly in resource-constrained environments like the Internet of Things (IoT). The Offline/Online (O/O) extension has gained attention for its effective signing process, as it allows for the precomputation of operations beforehand. Informally, in this kind of scheme, signing a message involves two stages. The first is an offline stage that requires moderate computation but can be done in advance, before the message is known. The second is an online stage that starts once the message is available, using the precomputed data to complete the signature more quickly. This survey provides the first comprehensive overview of the adaptation of O/O constructions for various resource-constrained use cases, such as IoT, low-power and mobile contexts, and other contexts, such as blockchains, where latency is a crucial metric. We reviewed 36 O/O DS schemes, focusing on trends, challenges, and applications. We examined the benefits, drawbacks, and potential areas for improvement of the reviewed O/O schemes, as well as identifying the features inherent to the O/O design. Our analysis indicates that the latter effectively mitigates issues related to expensive computations, delays, and high energy consumption, provided that the increased storage demands for offline signatures are manageable. This survey lays the groundwork for future research and development of O/O schemes, driven by the increasing prevalence of constrained environments where these schemes are particularly useful.

Related publication: [10]

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

EA AUDIT-PAIR

Participants: Héber Arcolezi, Claude Castelluccia, Mohamed Maouche, Antoinette Boutet, Mathieu Cunche, Clémentine Gritti.

- Title: AUDIT-PAIR: Algorithmic Auditing of Privacy and Fairness
- Website: <https://team.inria.fr/auditpair/en/>
- Type: Equipe Associée Inria
- Duration: 2024 - 2026
- Inria - PRIVATICS and UQAM et ÉTS Montréal
- Abstract: This Associated Team focusses on three topics: Privacy-Preserving Machine Learning, Differential Privacy Auditing, and Bias Detection and Mitigation in Machine Learning.

10.2 International research visitors

10.2.1 Visits of international scientists

Inria International Chair

Participants: Cristiana Santos.

Cristiana Santos - an Assistant Professor of Law from Utrecht University - has been an Inria International Chair during 2025, with whom we have collaborated on numerous projects, involving PRIVATICS members (Nataliia Bielova, Vincent Roca, Gilles Mertens), resulting in numerous joint publications and feedback to public consultation from the EU Commission and the EU Data Protection Board.

Participants: Nicolas Papernot.

Nicolas Papernot ([website](#)) is Assistant Professor at the University of Toronto and has been awarded an Inria International Chair in 2025, co-hosted by the PREMEDICAL and PRIVATICS Inria teams. His work lies at the intersection of security, privacy, and AI. Nicolas will work along with members of the IPoP project (PEPR Cybersécurité) and SSF-ML-DH (PEPR Santé Numérique).

10.3 European initiatives

10.3.1 Digital Europe

NoLeFa-84

- Title: NoLeFa-84
- Type: DIGITAL-CSA
- Duration: Nov 2024 - April 2027
- Coordinator: Inria
- Others partners: Inria PRIVATICS, LNE, Numalis, Piccadilly Labs, Leiwand AI.
- Abstract: The NoLeFa-84 project sets in the ambitious frame of the imminent entry into force of the AI Act, with gradual application of obligations over the next 3 years. The cornerstone of the AI Act application will be AI testing, which is a shared challenge for both the ecosystem and the authorities. The core activities of this project is to develop a suite of AI testing tools and procedures and contributing to AI Act standards on both the technical and the governance sides.

10.4 National initiatives

10.4.1 ANR

IPoP (PEPR Cybersecurity)

- Title: Interdisciplinary Project on Privacy
- Type: PEPR Cybersécurité / France 2030
- Duration: July 2022 - June 2028
- Coordinator: Inria - PRIVATICS, Antoine Boutet
- Others partners: Inria COMET / MAGNET / PETRUS / MULTISPEECH and SPIRALS teams, CNRS - DCS lab., INSA CVL - LIFO lab., Univ. Grenoble Alpes - 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 - PREMEDICAL
- Others partners: Inria PRIVATICS, Inria EPIONE, Inria MAGNET, 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.

GTTP

- Title: GTTP
- Type: ANR
- Duration: 2025 - 2028
- Coordinator: ENS Lyon
- Others partners: Inria - PRIVATICS, Sentiens
- Abstract: Due to the critical role of geolocalization in the ecosystem, performance requirements are high and diverse: tracking systems require accuracy, global coverage, indoor and outdoor, energy efficiency to support millions of IoT assets, real-time services, low latency to respond to customer requests, user-centric services, and strict privacy requirements. However, no existing geolocalization solution based Internet (IP geolocation), LPWAN networks or 5G , on satellites (GNSS) meets all these requirements simultaneously. They are unreliable, insecure, non-interoperable and non-scalable, because innovations in this area have remained siloed over the years. To address this issue, the GTTP project aims to study, design and prototype a protocol architecture capable of meeting the aforementioned requirements.

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 reimaged 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.

AI-PULSE

- Title: Aligning Privacy, Utility, and Fairness for Responsible AI
- Type: ANR JCJC grant
- Duration: 2025 - 2029
- Coordinator: Inria - PRIVATICS, Héber Hwang Arcolezi
- Abstract: In the era of Big Data, the development of AI systems presents both opportunities and challenges, particularly concerning privacy and fairness. AI-PULSE aims to address the interplay between Differential Privacy (DP) and fairness in Machine Learning (ML). While DP offers a promising framework for quantifying the privacy-utility trade-off, fairness in ML is essential for preventing discrimination. Although the current landscape between DP and fairness is multifaceted, this project seeks to explore how they can coexist harmoniously. Through a multidisciplinary approach, combining privacy, fairness, and utility metrics, AI-PULSE will provide novel insights into responsible AI. By investigating theoretical and practical aspects, it aims to move beyond central DP to local DP guarantees and to foster the adoption of privacy and fairness as fundamental practices in AI development. Our developed tools will be open-sourced to benefit both research and industry, especially those where AI systems for decision-making are prevalent. Moreover, through our collaboration with the French Data Protection Authority CNIL, we will ensure that the output of this project better suits their needs for advancing in Responsible AI.

RAIDAC+

- Title: Responsible AI: Design, Regulation, and Conformity
- Type: MIAI Cluster Chair (ANR 2030)
- Duration: 2025 - 2029
- Coordinator: Inria - PRIVATICS (Héber Hwang Arcolezi) & UGA - CESICE (Theodore Christakis)
- Abstract: The Responsible AI Chair (RAIDAC+) is a transdisciplinary initiative that integrates pioneering research in AI privacy, fairness, transparency, and robustness with cutting-edge expertise in AI regulation, law, and policy. As AI systems become ubiquitous—governing access to social services, guiding decisions in healthcare, informing public security strategies, and even facilitating novel neurotechnologies—they confront us with urgent ethical, legal, and societal challenges. Building on the solid foundations of two distinct yet complementary approaches—(1) the “Responsible AI” project focused on privacy risks, fairness, and explainability, and (2) the “Legal and Regulatory Implications of AI” Chair that addressed legal and governance frameworks (GDPR, EU AI Act, Law Enforcement Directive. . .)—RAIDAC+ uniquely unites computer scientists, legal scholars, and industry partners. Together, we aim to help data-driven innovation in Europe through ethically sound, legally compliant, and socially beneficial AI technologies.

FONDUE

- Title: FONDUE: artiFicial intelligence fOR administrative aND bUusiness procEsses
- Type: MIAI Cluster Chair (ANR 2030)
- Duration: 2025 - 2029
- Coordinator: Université Savoie Mont Blanc, Annecy
- Abstract: Artificial intelligence (AI), like large language models (LLMs) and expert systems, is transforming business and administrative processes, impacting information retrieval, regulation/rule matching, and decision-making. AI has a significant societal impact but also poses challenges, including bias, ethical concerns, and the risk of reinforcing societal inequalities. Business AI systems must integrate monitoring components from the design phase to assess ethical, social, and regulatory impacts alongside operational metrics. The FONDUE project aims to apply AI methodologies to improve all stages of business processes while addressing bias and ethical risks. It emphasizes integrating monitoring and evaluation components into AI systems to enhance transparency, explainability, and societal benefits.

10.4.2 Inria Exploratory Action (AEx)

DATA4US (Personal Data Transparency for web USERS)

- Participants: Cedric Lauradoux, Nataliia Bielova
- Duration: 2020-2025
- 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.

10.4.3 Others

INTERFERE

- Title: INTERFERE: Stressing Systems Security Through on the Fly Network Traffic Generation
- Type: FIL (Inter-laboratory projects)
- Duration: 2024 - 2025
- Coordinator: Inria MiMove
- Others partners: Inria PRIVATICS
- Abstract: The INTERFERE project focuses on generating sequential data such as network traffic or query streams to stress systems facing security challenges. This includes 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 while maintaining control over the events represented. By providing research into auditing tools, this project plays a key role in the development of new methods and technologies to strengthen system and data security.

10.5 Public policy support

CNIL (French Data Protection Authority)

- Participants: Claude Castelluccia
- Claude Castelluccia is appointed member of the CNIL Board (“commissaire CNIL”) as qualified public member for his expertise on digital sciences and privacy questions, for the 2021-2026 period. As such he is in charge of several domains and contributes to the doctrine of the French Data Protection Authority.

PANAME (Privacy Auditing of AI Models)

- Coordinator: CNIL
- Others partners: PEReN, ANSSI
- Participants: Antoine Boutet, Mohamed Maouche, Heber Hwang Arcolezi
- Duration: 2025-2026
- Abstract: PANAME aims to develop a tool for auditing the privacy of AI models. More specifically, the objective is to develop a software library enabling the efficient and cost-effective implementation of certain technical privacy assessment tests that stakeholders in the AI ecosystem may need to perform to evaluate the GDPR compliance of an AI model.

11 Dissemination

Participants: Héber Arcolezi, Nataliia Bielova, Antoine Boutet, Claude Castelluccia, Mathieu Cunche, Clémentine Gritti, Cédric Lauradoux, Mohamed Maouche, Vincent Roca.

11.1 Promoting scientific activities

11.1.1 Scientific events: organisation

- Privacy Alpine Seminar, March 2025, Corrençon en Vercors (Antoine Boutet, Clémentine Gritti)
- Dagstuhl Seminar on Online Privacy: Transparency, Advertising, and Dark Patterns, January 2025 (Nataliia Bielova)
- Panel "Can a fair power balance be achieved in the Web ecosystem with the help of Computer Science research?" at Computers, Privacy and Data Protection (CPDP), May 2025 (Nataliia Bielova)
- GDR RSD / ASF Winter School on Distributed Systems and Networks, February 2025 (Antoine Boutet)
- IPoP Workshop "Créer la confiance dans les données", PEPR Cyberday - European Cyber Week, Rennes, November 2025. IPoP Workshop "Créer la confiance dans les données", PEPR Cyberday - European Cyber Week, Rennes, November 2025 (Antoine Boutet)

Member of the organizing committees

- Mathieu Cunche and Nataliia Bielova are members of the steering committee of APVP (Atelier sur la Protection de la Vie Privée)
- Nataliia Bielova, Claude Castelluccia, Mathieu Cunche are jury members of the CNIL-Inria Privacy Award.

11.1.2 Scientific events: selection

Member of the conference program committees

- ACISP 2025, Clémentine Gritti
- INSCRYPT 2025, Clémentine Gritti
- ACM SIGAPP SAC 2025, Clémentine Gritti
- ACM WiSec 2025, Mathieu Cunche
- DPM 2025, Mathieu Cunche
- ACM CCS 2025, Héber Hwang Arcolezi
- PETS 2025, Héber Hwang Arcolezi
- USENIX Security 2025, Héber Hwang Arcolezi
- ICLR 2025, Héber Hwang Arcolezi
- ConPro 2025, Nataliia Bielova
- PETS 2025, Nataliia Bielova
- ACM CCS 2025, Nataliia Bielova
- ACM TheWebConf 2025, Antoine Boutet

Reviewer

- UAI 2025, Mohamed Maouche

11.1.3 Journal

Member of the editorial boards

- IACR Journal of Cryptology, Clémentine Gritti

Reviewer - reviewing activities

- IEEE Transactions on Information Forensics and Security (TIFS), Mohamed Maouche, Héber Hwang Arcolezi, Clémentine Gritti
- IEEE Transactions on Dependable and Secure Computing (TDSC), Mohamed Maouche, Héber Hwang Arcolezi
- ACM Transactions on Privacy and Security (TOPS), Héber Hwang Arcolezi
- Computers & Security, Héber Hwang Arcolezi

11.1.4 Invited talks

- Anonymization by Design of Language Modeling (Bernoulli Lab Seminar, Paris, March 2025), Antoine Boutet
- Protection of Federated Learning Against Inference Attacks (Scientific Day on Federated and Decentralized Learning, FIL, Lyon, June 2025), Antoine Boutet
- Towards the Anonymization of the Language Modeling (EPFL Seminar, Lausanne, May 2025), Antoine Boutet
- L'interdisciplinarité au service de la cybersécurité: présentation du projet IPoP (Rencontres Sécurité Informatique et Sciences Humaines et Sociales, Paris, January 2025), Antoine Boutet
- Challenges of decentralized AI Privacy and Robustness (Unite! school, 04/11/25), Mohamed Maouche
- Interplay of privacy and fairness in machine learning (Unite! school, 04/11/25), Héber Hwang Arcolezi
- Intersections of Fairness and Privacy: A Local Differential Privacy Perspective (9th GDR RSD/ASF Winter School on Distributed Systems & Networks 2025, 06/02/25), Héber Hwang Arcolezi
- The EU's Next Move on Online Advertising (Computers, Privacy and Data Protection, CPDP.ai 2025, Brussels, Belgium, May 2025), Nataliia Bielova
- Regulating Consent and Dark Patterns on the Web: A Transdisciplinary Approach with Web Measurements, Law, and HCI (Distinguished lecture at the Cyber Security in the Age of Large-Scale Adversaries (CASA) Cluster of Excellence, Bochum, Germany, April 2025), Nataliia Bielova
- Privacy regulations and the Web (World Wide Web Consortium (W3C) Advisory Committee event, Sophia Antipolis, April 2025), Nataliia Bielova
- Challenges in compliance auditing of consent management at scale (Scoping Emerging Consent Practices in the Digital Ecosystem virtual roundtable of the Organisation for Economic Co-operation and Development (OECD), March 2025), Nataliia Bielova
- Regulating Consent and Dark Patterns on the Web: A Transdisciplinary Approach with Web Measurements, Law, and HCI (Oslo Metropolitan University, November 2025), Nataliia Bielova

11.1.5 Leadership within the scientific community

- Co-chair of the Working Group on privacy protection (GT-PVP) of the GDR Sécurité, Mathieu Cunche

11.1.6 Scientific expertise

- ANR Comité TSIA Numérique et Mathématiques, Mohamed Maouche.
- Mathieu Cunche is a member of the scientific advisory board of the *13 Novembre* program.
- Vincent Roca was reviewer/scientific expert for the ANR call for projects, March 2025.

11.1.7 Research administration

- 'Chargé de communication' for the Lyon Computer Science Federation (FIL), Mohamed Maouche.
- Member of Comité de Centre of Inria centre of University Cote d'Azur, Nataliia Bielova
- Treasurer and then president of ASF (ACM Sigops France), Antoine Boutet

11.2 Teaching - Supervision - Juries - Educational and pedagogical outreach

11.2.1 Juries

- Antoine Boutet was a rapporteur for the PhD thesis of Jade Garcia Bourrée, Inria Rennes, October 2025.
- Antoine Boutet was a rapporteur for the PhD thesis of Pierre-Marie Lechevalier, IMT Atlantique, June 2025.
- Antoine Boutet was a reviewer for the PhD thesis of Antoine Barczewski, Inria Lille, October 2025.
- Antoine Boutet was a reviewer for the PhD thesis of Rezak Aziz, CNAM, December 2025.
- Clémentine Gritti was a rapporteur for the PhD thesis of Lise Millerjord, NTNU (Norway), September 2025.
- Mathieu Cunche was a rapporteur for the PhD thesis of Soumaya Boussha Sorbonne Université / Eurecom, December 2025.
- Mathieu Cunche was a reviewer for the PhD thesis of Aniketh Girish, Universidad Carlos III de Madrid, June 2025.
- Mathieu Cunche was a rapporteur for the HDR of Daniele Antonioli, Institut Polytechnique Paris, Ecole Polytechnique and EURECOM, June 2025.
- Nataliia Bielova was a rapporteur for the PhD thesis of Simon Koch, echnical University Braunschweig (Germany), April 2025.
- Vincent Roca was a rapporteur for the HDR of Pierre Laperdrix, "Untangling the Web TRacKing Ecosystem to Design Effective Defenses", Université de Lille, October 2025.
- Vincent Roca was a rapporteur for the PhD thesis of Maxime Huyghe, "Exploration automatique de l'impact des paramètres de configuration sur les empreintes de navigateurs", Université de Lille, November 2025.

11.2.2 Educational and pedagogical outreach

Most of the PRIVATICS members' lectures are given at INSA-Lyon (Antoine Boutet, Mathieu Cunche and Clementine Gritti are professor/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 law students.

As part of the Inria CHICHE program, Cédric Lauradoux contributes to more than half of the Grenoble research center CHICHE. He also participates in "Fête de la Science", in the MathC2+ week, and in Math Olympiads.

Cédric Lauradoux has been appointed CHICHE Project Manager for the Centre Inria de l'Université Grenoble Alpes.

11.3 Popularization

11.3.1 Others science outreach relevant activities

- Nataliia Bielova has presented her work on privacy issues on the Web at the "Fête de la Science" (Scientific Faire) in Juan les Pins, France, October 2025.
- Antoine Boutet has led a workshop (Inspecter ce que votre historique de mobilité révèle sur vous) during the "Fête de la science", Insa-Lyon, France, October 2025.

12 Scientific production

12.1 Major publications

- [1] G. P. Kancherla, N. Bielova, C. Santos and A. Bichhawat. 'Johnny Can't Revoke Consent Either: Measuring Compliance of Consent Revocation on the Web'. In: PETS 2025 - Privacy Enhancing Technologies Symposium. Vol. 2025. 4. Washington / Virtual, United States, 14th July 2025, pp. 329–347. DOI: [10.56553/popets-2025-0133](https://doi.org/10.56553/popets-2025-0133). URL: <https://hal.science/hal-05474344>.
- [2] G. Mertens, N. Bielova, V. Roca and C. Santos. 'You Can't Trust Your Tag Neither: Privacy Leaks and Potential Legal Violations within the Google Tag Manager'. In: EuroS&P 2025 - 10th IEEE European Symposium on Security and Privacy. Venice (Ca' Foscari University), Italy, 2025, pp. 1–20. URL: <https://hal.science/hal-05032798>.
- [3] A. K. Mishra, G. Gagnon, M. Cunche and S. Gambs. 'QRisk: Think Before You Scan QR codes'. In: *Lecture Notes in Computer Science (LCNS)*. ARES 2025 - 20th International Conference on Availability, Reliability and Security. Ghent (BE), Belgium, 2025, pp. 1–22. URL: <https://hal.science/hal-04987069>.
- [4] J. Nicolas, C. Sabater, M. Maouche, M. Coates and S. B. Mokhtar. 'PriviRec: Confidential and Decentralized Graph Filtering for Recommender Systems'. In: CIKM 2025 - ACM International Conference on Information and Knowledge Management. Séoul, South Korea: ACP, 2025, pp. 1–9. DOI: [10.1145/3746252.3761152](https://doi.org/10.1145/3746252.3761152). URL: <https://hal.science/hal-05308553>.
- [5] O. Touat, J. Brunon, Y. Belal, J. Nicolas, C. Sabater, M. Maouche and S. Ben Mokhtar. 'Exposing the Vulnerability of Decentralized Learning to Membership Inference Attacks Through the Lens of Graph Mixing'. In: MIDDLEWARE '25: Proceedings of the 26th International Middleware Conference. Nashville, TN, United States: ACM, 15th Dec. 2025, pp. 180–194. DOI: [10.1145/3721462.3770770](https://doi.org/10.1145/3721462.3770770). URL: <https://hal.science/hal-04933985>.

12.2 Publications of the year

International journals

- [6] H. H. Arcolezi, K. Makhoul and C. Palamidessi. 'Group fairness under obfuscated sensitive information'. In: *Journal of Computer Security* (14th May 2025), pp. 1–26. DOI: [10.1177/0926227X251330212](https://doi.org/10.1177/0926227X251330212). URL: <https://inria.hal.science/hal-05070800> (cit. on p. 13).
- [7] J. Gunawan, C. M. Gray, C. Santos and N. Bielova. 'Leveraging interdisciplinary methods for evidence collection in enforcement: Dark patterns as a case study'. In: *Internet Policy Review* 14.4 (18th Nov. 2025). DOI: [10.14763/2025.4.2047](https://doi.org/10.14763/2025.4.2047). URL: <https://hal.science/hal-05473920> (cit. on p. 20).

- [8] G. P. Kancherla, N. Bielova, C. Santos and A. Bichhawat. ‘Johnny Can’t Revoke Consent Either: Measuring Compliance of Consent Revocation on the Web’. In: *Proceedings on Privacy Enhancing Technologies* 2025.4 (14th July 2025), pp. 329–347. DOI: [10.56553/popets-2025-0133](https://doi.org/10.56553/popets-2025-0133). URL: <https://hal.science/hal-05474344> (cit. on p. 20).
- [9] S. Péliissier, A. K. Mishra, M. Cunche, V. Roca and D. Donsez. ‘Efficiently linking LoRaWAN identifiers through multi-domain fingerprinting’. In: *Pervasive and Mobile Computing* 112 (Aug. 2025), p. 102082. DOI: [10.1016/j.pmcj.2025.102082](https://doi.org/10.1016/j.pmcj.2025.102082). URL: <https://inria.hal.science/hal-05120767> (cit. on p. 18).
- [10] J. Samandari and C. Gritti. ‘Online/Offline Digital Signatures: A Systematic Literature Review’. In: *IEEE Access* 13 (2025), pp. 90991–91011. DOI: [10.1109/ACCESS.2025.3570689](https://doi.org/10.1109/ACCESS.2025.3570689). URL: <https://hal.science/hal-05420540> (cit. on p. 23).
- [11] C. Santos, C. M. Gray, N. Bielova and S. Ahuja. ‘Usable and Lawful: Can Consent Be Both?’ In: *Information & Communications Technology Law* (29th Dec. 2025). DOI: [10.1080/13600834.2025.2610581](https://doi.org/10.1080/13600834.2025.2610581). URL: <https://inria.hal.science/hal-05304243> (cit. on p. 21).

International peer-reviewed conferences

- [12] S. Ahuja, J. Gunawan, N. Bielova and C. T. Santos. ‘Towards Key Contributing Factors in Identifying Dark Pattern Autonomy Violations under the EU Digital Services Act’. In: *ACM Digital library. DIS 2025 - ACM Designing Interactive Systems Conference. DIS ’25 Companion: Companion Publication of the 2025 ACM Designing Interactive Systems Conference*. Funchal, Portugal: ACM, 5th July 2025, pp. 501–507. DOI: [10.1145/3715668.3736336](https://doi.org/10.1145/3715668.3736336). URL: <https://inria.hal.science/hal-05304453> (cit. on p. 21).
- [13] A. Akram, H. N. H. Pham, M. Önen and C. Gritti. ‘SAAFL: Secure Aggregation for Label-Aware Federated Learning’. In: *IFIP SEC 2025 - 40th International Conference on ICT Systems Security and Privacy Protection*. Maribor, Slovenia, 2025, pp. 1–14. URL: <https://hal.science/hal-05066911> (cit. on p. 13).
- [14] H. H. Arcolezi, M. Alishahi, A.-A. Bendoukha and N. Kaaniche. ‘Fair play for individuals, foul play for groups? Auditing anonymization’s impact on ML fairness’. In: *ECAI 2025 : 28th European Conference on Artificial Intelligence*. 28th European Conference on Artificial Intelligence (ECAI). Vol. 413. Frontiers in Artificial Intelligence and Applications. Bologna, Italy: IOS Press, 21st Oct. 2025, pp. 1009–1018. DOI: [10.3233/FAIA250909](https://doi.org/10.3233/FAIA250909). URL: <https://inria.hal.science/hal-05386310> (cit. on p. 14).
- [15] Y. Belal, M. Maouche, S. B. Mokhtar and A. Simonet-Boulogne. ‘Inferring Communities of Interest in Collaborative Learning-based Recommender Systems’. In: *International Conference on Distributed Computing Systems*. 45th IEEE International Conference on Distributed Computing Systems (ICDCS 2025). 2025 IEEE 45th International Conference on Distributed Computing Systems (ICDCS). Glasgow, United Kingdom: IEEE, 2025. DOI: [10.1109/ICDCS63083.2025.00056](https://doi.org/10.1109/ICDCS63083.2025.00056). URL: <https://hal.science/hal-05007813> (cit. on p. 14).
- [16] A.-A. Bendoukha, H. H. Arcolezi, N. Kaaniche, A. Boudguiga, R. Sirdey and P.-E. Clet. ‘FADE: federated aggregation with discrimination elimination’. In: *FAccT 2025: ACM Conference on Fairness, Accountability, and Transparency*. FAccT 2025 - ACM Conference on Fairness, Accountability, and Transparency (FAccT). Athènes, Greece, 2025, pp. 1–14. URL: <https://hal.science/hal-05105146> (cit. on p. 15).
- [17] A. Boutet and V. Morel. ‘“I’m Not for Sale” -Perceptions and Limited Awareness of Privacy Risks by Digital Natives About Location Data’. In: *ICWSM 2025 - International AAAI Conference on Web and Social Media*. Copenhagen, Denmark, 2025, pp. 1–13. URL: <https://hal.science/hal-05041457> (cit. on p. 21).
- [18] T. Curelariu. ‘The Impact of EU Digital Regulation on the Protection of Vulnerability Disclosure Researchers’. In: *EU Digital Technologies and Policy Conference, Abstracts and Contributions*. EU Digital Technologies and Policy Conference (EUDTP 2025). Brussels, Belgium, 1st June 2025. URL: <https://hal.science/hal-05481821> (cit. on p. 23).

- [19] A. Losty, A. K. Mishra, M. Cunche and A. M. Mandalari. ‘Towards Operational and Security Best Practices for DNS in the Internet of Things’. In: ANRW 2025 - Applied Networking Research Workshop. Madrid, Spain, 2025, pp. 1–4. DOI: [10.1145/3744200.3744764](https://hal.science/hal-05110445). URL: <https://hal.science/hal-05110445> (cit. on p. 18).
- [20] G. Mertens, N. Bielova, V. Roca and C. Santos. ‘You Can’t Trust Your Tag Neither: Privacy Leaks and Potential Legal Violations within the Google Tag Manager’. In: EuroS&P 2025 - 10th IEEE European Symposium on Security and Privacy. Venice (Ca’ Foscari University), Italy, 2025, pp. 1–20. URL: <https://hal.science/hal-05032798> (cit. on p. 18).
- [21] A. K. Mishra and M. Cunche. ‘StateFi: Effectively Identifying Wi-Fi Devices through State Transitions’. In: *StateFi: Effectively Identifying Wi-Fi Devices through State Transitions*. WISEC 2026 - 19th ACM Conference on Security and Privacy in Wireless and Mobile Networks. Saarbrücken, Germany, 30th June 2026. URL: <https://hal.science/hal-05487604>.
- [22] A. K. Mishra and M. Fiore. ‘k-scale: k-Anonymizing Millions of Trajectories’. In: *IEEE INFOCOM 2026 Conference Proceedings*. IEEE INFOCOM 2026. Tokyo, Japan, 18th May 2026. URL: <https://hal.science/hal-05407897> (cit. on p. 19).
- [23] A. K. Mishra, G. Gagnon, M. Cunche and S. Gams. ‘QRisk: Think Before You Scan QR codes’. In: *Lecture Notes in Computer Science (LCNS)*. ARES 2025 - 20th International Conference on Availability, Reliability and Security. Ghent (BE), Belgium, 2025, pp. 1–22. URL: <https://hal.science/hal-04987069> (cit. on p. 19).
- [24] J. Nicolas, C. Sabater, M. Maouche, M. Coates and S. B. Mokhtar. ‘PriviRec: Confidential and Decentralized Graph Filtering for Recommender Systems’. In: CIKM 2025 - ACM International Conference on Information and Knowledge Management. Séoul, South Korea: ACP, 2025, pp. 1–9. DOI: [10.1145/3746252.3761152](https://hal.science/hal-05308553). URL: <https://hal.science/hal-05308553> (cit. on p. 15).
- [25] R. Taiello, C. Gritti, M. Önen and M. Lorenzi. ‘Buffalo: A Practical Secure Aggregation Protocol for Buffered Asynchronous Federated Learning’. In: CODASPY 2025 - 15th ACM Conference on Data and Application Security and Privacy. Pittsburgh, United States: ACM, 2025, pp. 1–12. DOI: [10.1145/3714393.3726498](https://hal.science/hal-05034678). URL: <https://hal.science/hal-05034678> (cit. on p. 15).
- [26] H. Zhang, A. K. Mishra and H. H. Arcolezi. ‘Demo: Exploring Utility and Attackability Trade-offs in Local Differential Privacy’. In: *CCS ’25: Proceedings of the 2025 ACM SIGSAC Conference on Computer and Communications Security*. CCS 2025 - ACM SIGSAC Conference on Computer and Communications Security. Taipei, Taiwan: ACM, 19th Nov. 2025, pp. 4728–4730. DOI: [10.1145/3719027.3760706](https://inria.hal.science/hal-05386311). URL: <https://inria.hal.science/hal-05386311> (cit. on p. 16).

Conferences without proceedings

- [27] M. Bernelin and A. Boutet. ‘PIA : Enseigner la protection des données personnelles dans l’interdisciplinarité’. In: RESSI 2025 - Rendez-vous de la Recherche et de l’Enseignement de la Sécurité des Systèmes d’Information. Quimper, France, 2025, pp. 1–5. URL: <https://hal.science/hal-05041374> (cit. on p. 23).
- [28] Y. Li, C. Eichler, N. Anciaux, H. H. Arcolezi and J. M. de Fuentes. ‘bench-MIA: Towards automatic multi-lever benchmark construction for MIA evaluation’. In: APVP 2025 - 15ème Atelier sur la protection de la vie privée. Chasseneuil-du-Poitou, France, 2025. URL: <https://inria.hal.science/hal-05287192> (cit. on p. 16).
- [29] O. Touat, J. Brunon, Y. Belal, J. Nicolas, C. Sabater, M. Maouche and S. Ben Mokhtar. ‘Exposing the Vulnerability of Decentralized Learning to Membership Inference Attacks Through the Lens of Graph Mixing’. In: MIDDLEWARE ’25: Proceedings of the 26th International Middleware Conference. Nashville, TN, United States: ACM, 15th Dec. 2025, pp. 180–194. DOI: [10.1145/3721462.3770770](https://hal.science/hal-04933985). URL: <https://hal.science/hal-04933985> (cit. on p. 17).

Reports & preprints

- [30] S. Ahuja, J. Gunawan, N. Bielova and C. T. Santos. *Dark Patterns and the EU Digital Services Act: Mapping Autonomy Violations and Design Factors*. 2025. URL: <https://hal.science/hal-05301214>.
- [31] S. Ahuja, G. P. Kancherla, C. T. Santos, N. Bielova and A. Bichhawat. *How Usable is Consent Withdrawal on the Web? UI Requirements and Expert Evaluation*. 2025. URL: <https://hal.science/hal-05302086>.
- [32] Y. Belal, M. Maouche, S. B. Mokhtar and A. Simonet-Boulogne. *GRANITE : a Byzantine-Resilient Dynamic Gossip Learning Framework*. 2025. URL: <https://hal.science/hal-05059491> (cit. on p. 17).
- [33] N. Bielova, C. Santos, C. M. Gray, A. Rossi, B. Schaffner, A. Stoeber, H. Krahl and J. Gunawan. *Feedback to the EU Commission’s Call for evidence for an impact assessment — Ares(2025)5829481 on the Digital Fairness Act*. Inria & Université Cote d’Azur, Sophia Antipolis, France, 24th Oct. 2025, pp. 1–22. URL: <https://hal.science/hal-05330366> (cit. on p. 22).
- [34] A. Boutet and L. Magnana. *Leverage Unlearning to Sanitize LLMs*. Inria Lyon, 1st Sept. 2025. URL: <https://hal.science/hal-05485551>.
- [35] G. Mertens, N. Bielova, V. Roca, A. Bouhoula and M. Akassab. *An Analysis of Client-and Server-Side Google Tag Manager and its Tags on the Web*. 6th Feb. 2026. URL: <https://hal.science/hal-05466083>.
- [36] J. Sénéchal, A. Boutet, L. Magnana and H. Zimmermann. *Towards the Anonymization of the Language Modeling*. 4th May 2025. DOI: [10.48550/arXiv.2501.02407](https://doi.org/10.48550/arXiv.2501.02407). URL: <https://lilloa.hal.science/hal-05223726> (cit. on p. 17).
- [37] L. Zard, O. Goga, A. El fraihi and N. Bielova. *Feedback to the European Data Protection Board’s Guidelines 3/2025 on the interplay between the DSA and the GDPR (Version 1.1) - Advertisement*. Inria & Université Cote d’Azur, Sophia Antipolis, France, 31st Oct. 2025, pp. 1–20. URL: <https://hal.science/hal-05387881> (cit. on p. 22).

Other scientific publications

- [38] A. Akram, C. Gritti and M. Önen. ‘Privacy-Preserving Federated Learning’. In: Eurecom presentation 2025. Sophia Antipolis, France, 2025. URL: <https://hal.science/hal-05127285>.

Scientific popularization

- [39] C. Adam and C. Lauradoux. ‘Feedback from a debate game to raise awareness about the use of AI for sanitary surveillance’. In: *Recherches et recherches-actions en didactique de l’informatique 2* (Aug. 2025), pp. 1–20. URL: <https://hal.science/hal-05374049> (cit. on p. 19).