PRIVGEN

Privacy-preserving sharing and processing of genetic data

LaTIM Inserm UMR 1101
Ascola, LINA UMRS CNRS 6241
Inserm UMR 1078

in collaboration with

Labex Genmed

Partners

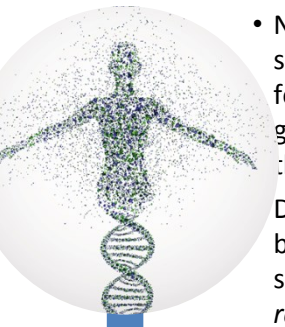
Context

- **Cloud Computing and data and applications Outsourcing**

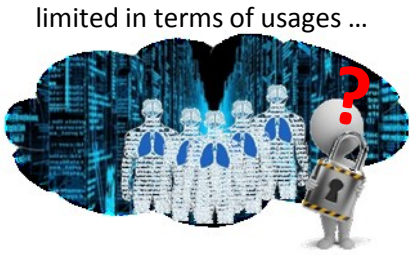
  - A successful paradigm to flexibility store, share and process large amount of data while minimizing costs
  - Targets: individuals, companies, governments
  - Domains: all you can imagine (health, social networks …)

- **Security needs of outsourced applications and data are worsened**
  - Owners loss the control on their data and applications (***confidentiality***, ***integrity***, ***availability***?)
  - Service provider may in turn transmit data to third-party service providers (***traceability***, ***intellectual/scientific ownership protection***?)
  - Storage by the service providers of data issued from different sources (***privacy***?)

- **Sharing of outsourced genetic data and applications – more than an experimental framework**



  - Needs for national and international sharing of genetic data of individuals for better decryption of the human genome so as to improve diagnosis, therapy …
  - Data highly personal and of dynamic behavior, covering a large security spectrum needs (*privacy*, *data reliability – integrity + authenticity* -, *scientific ownership* …)
  - Distributed applications
  - Different initiatives (e.g. beacons) with identified security weaknesses , so limited in terms of usages …



Objectives

- **Provide a response to the limitations of actual security solutions in the cloud**
  - Cloud applications have to satisfy a large number of different security and privacy properties at once ➔ Requires to make interacting different security mechanisms
  - Cloud applications involve computations at different sites that are executed on behalf of multiple stakeholders
- **Two research axis**
  - Means for the composition of several security and privacy mechanisms applied to compositions of complex computations in the cloud
  - New multipurpose security mechanisms able to satisfy simultaneously several security objectives at once (e.g. confidentiality, privacy, traceability)

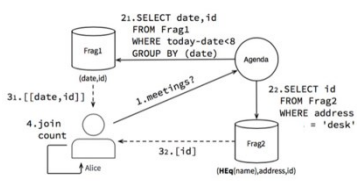
Challenge 1 - Mechanisms for a continuous digital content protection

- **Objective:** Merging of different security mechanisms into one adaptable and configurable digital content protection tool for continuous and multipurpose security protection

  - *Contribution 1* – Identify the constraints and limits of actual security tools and of their simple juxtaposition when composing complex computations in outsourced environments
  - *Contribution 2* – Achieve a continuous protection of shared and mutualized data, by means of joint-security mechanisms (watermarking, encryption, fragmentation …) compliant and configurable by a security service composition language

Challenge 2 - Composition of security and privacy-protection mechanisms

- **Objective:** Provide a development approach for secure and privacy-preserving distributed genetic applications



  - *Contribution 1* – A composition theory for watermarking, cryptographic and fragmentation-based security and privacy properties.
  - *Contribution 2* – Programming support that satisfies security and privacy properties by construction and formal verification.

Challenge 3 - Distributed processing and sharing of genetic data

- **Objective:** Provide a method and platform for sharing a minimum set of relevant genomic information while maintaining privacy

  - *Contribution 1* – Methods and chniques for the sharing of genetic ata under strong privacy properties nstraints, harnessing results from allenges 1 and 2
  - *Contribution 2* – A platform supporting the distributed execution of applications over shared genetic data

People

- **G. Coatrieux**, Pr., LaTIM Inserm UMR 1101, Télécom Bretagne
- **M. Südholt**, Pr., Ascola, LINA UMR CNRS 6241, Mines de Nantes
- **E. Genin**, Dr., Inserm UMR 1078
- **J-F. Deleuze**, Dr., CNG
- **V. Meyer**, IR, CNG

- **D. Niyitegeka**, Ph.D Student, LaTIM Inserm UMR 1101, Télécom Bretagne
- **F. Boujdad**, Ph.D Student, Ascola, LINA UMR CNRS 6241, Mines de Nantes
- **D. Bouslimi**, IR, LaTIM Inserm UMR 1101, Télécom Bretagne
- **M. W. Pan**, IR, LaTIM Inserm UMR 1101, Télécom Bretagne

Selected publications

- R-A. Cherrueau, R. Douence, and M. Südholt. ***A Language for the Composition of Privacy-Enforcement Techniques***. In the 2015 IEEE Int. Symp. RATSP, Aug. 2015.
- R-A. Cherrueau, M. Südholt, and O. Chebaro. ***Adapting workflows using generic schemas: application to the security of business processes***. In 5th IEEE Int. Conf. on Cloud Technology and Science , p. 6., Dec. 2013.
- B. De Fraine, E. Ernst, and M. Südholt. ***Essential aop: The a calculus***. ACM Trans. Program. Lang. Syst., 34(3):12:1–12:43, Nov. 2012.
- D. Bouslimi, G. Coatrieux, M. Cozic, C. Roux. ***Data hiding in encrypted images based on predefined watermark embedding before encryption process***. Sig. Proc.: Image Comm. 47: 263-270 (2016)
- D. Bouslimi and G. Coatrieux. ***Encryption and Watermarking for Medical Image Protection***, chapter Medical Data Privacy Handbook. Springer, 2016.
- J. Franco-Contreras and G. Coatrieux. ***Robust watermarking of relational databases with ontology-guided distortion control***. IEEE Trans. Information Forensics and Security, 10(9):1939–1952, (2015)
- A. Saint Pierre and E. Génin. ***How important are rare variants in common disease?***. Briefings in Functional Genomics, 07, 2014.
- S. Gazal, M. Sahbatou, M-C. Babron, E. Génin Emmanuelle, and A-L. Leutenegger. ***Fsuite: exploiting inbreeding in dense snp chip and exome data***. Bioinformatics, 03 2014.
- M-C. Babron, M. De Tayrac, D.N. Rutledge, E. Zeggini, and E. Génin. ***Rare and low frequency variant stratification in the uk population: Description and impact on association tests***. PLoS ONE, 7(10), 10, (2012)