Multi-Owner data Sharing for Analytics and Integration respecting Confidentiality and Owner control

Reporting

Project Information

**MOSAICrOWN**
Grant agreement ID: 825333

**DOI**
10.3030/825333

Funded under
H2020-EU.2.1.1.

Overall budget
€ 3 203 750

EU contribution
€ 3 203 750

**Coordinated by**
UNIVERSITA DEGLI STUDI DI MILANO
Italy

Start date
1 January 2019

End date
31 December 2021

Closed project

Periodic Reporting for period 2 - MOSAICrOWN (Multi-Owner data Sharing for Analytics and Integration respecting Confidentiality and Owner control)

**Reporting period:** 2020-07-01 to 2021-12-31

**Summary of the context and overall objectives of the project**

The application of data analysis techniques over large data collections provides great benefits, from personal, to business, research, and social domains. The continuous evolution of ICT is enabling the realization of such data-driven innovations at a fast pace, towards the realization of digital data
markets supporting collaborative data sharing and analytics. A clear obstacle towards the realization of such potential and vision is represented by security and privacy concerns. Indeed, the (actual or perceived) loss of control over data and potential compromise of their confidentiality, also protected by privacy regulations, can have a strong detrimental impact on the realization of an open framework for enabling the sharing of information from multiple independent data owners.

The goal of MOSAICrOWN is to enable data sharing and collaborative analytics in multi-owner scenarios in a privacy-preserving way, ensuring proper protection of private/sensitive/confidential information. MOSAICrOWN has provided effective and deployable solutions allowing data owners to maintain control on the data sharing process, enabling selective and sanitized disclosure providing for efficient and scalable privacy-aware collaborative computations.

The practical objectives that MOSAICrOWN has pursued are as follows.

Objective 1 – Rich support of requirements, considering different aspects of protection and needs from different parties (data owners, as well as data subjects and privacy regulations) and addressing their satisfaction under different scenarios and threat models.

Objective 2 – Data governance framework, empowering owners with control on their data, enabling them to specify policies regulating protection of information and its selective disclosure in collaborative data platforms.

Objective 3 – Data wrapping, for supporting selective release, storage and analytics on data in the collaborative platform, while preventing (or limiting) access to the actual data content by other parties.

Objective 4 – Data sanitization, for enforcing privacy/confidentiality restrictions by producing information for the data market, or within the data market, while protecting the precise values in the original data sources.

Objective 5 – Effective exploitation, in real operational environments, demonstrating the applicability and flexibility of the project’s innovations and actual impact.

MOSAICrOWN has met all the objectives above by considering use cases providing rich and comprehensive requirements corresponding to real problems and market strategies of major players.

Work performed from the beginning of the project to the end of the period covered by the report and main results achieved so far

The project has advanced state of the art and produced innovations on all the planned objectives.

Objective 1 – Rich support of requirements. The project has considered three use cases, corresponding to real-world problems of the industrial partners. Thanks to the richness and complementarity of the use cases, the project has provided a comprehensive list of requirements to be addressed under different scenarios. The requirement analysis has covered different aspects related
to the effective protection of data in the digital data market, from basic storage, to fine-grained retrieval, and controlled sharing.

Objective 2 - Data governance framework. The project has defined the overall architecture of the data governance framework and of the policies regulating its behavior. The project has also produced a policy model and language enabling data owners to specify - and have enforced - policies on their data ingested, stored, and processed in the data market. The project has also implemented the policy engine, for the correct enforcement of access and usage restrictions to be enforced on data ingested in the data market.

Objective 3 - Data wrapping. The project has designed novel solutions for protecting data stored in the data market, also with consideration of distributed settings. Some of the proposed solutions have also been made available open source. The solutions included advanced approaches for strong resource protection and resource fragmentation with decentralized allocation, fine-grained access to wrapped data stored in the digital data market, execution of collaborative computations while ensuring full respect of the restrictions imposed by the authorizations.

Objective 4 - Data sanitization. The project has investigated privacy and utility metrics relevant for digital data market scenarios and has developed sanitization solutions considering both syntactic and semantic approaches. The solutions developed enable data owners and analysts to efficiently anonymize large data collections and sanitize data collected from multiple data owners for collaborative analytics, also enabling the assessment against membership attacks and guiding the parametrization of differential privacy. The solutions enabled also the distribution of anonymization work to different workers in a distributed system, hence providing efficient and effective anonymization of very large data collections.

Objective 5 - Effective exploitation. Industrial partners have designed and pursued exploitation presenting MOSAICrOWN and some findings to customers, at industrial events and at meetings. MOSAICrOWN results are also used by partners to enhance their internal research and product development. SAP has worked on additional anonymization techniques (to extend HANA’s capabilities), and privacy interpretations for machine learning in business applications; EISI has integrated MOSAICrOWN techniques into data management products within EISI technology portfolio and EISI has presented the project to a number of internal executive meetings and external industry groups within Ireland; MC has refined its data security practices and has investigated the integration of MOSAICrOWN techniques into its day-to-day operations and offerings. In addition, solutions developed by academic partners have been made available open-source to the wide research and development community, enabling others to build on the project results.

Progress beyond the state of the art and expected potential impact (including the socio-economic impact and the wider societal implications of the project so far)
MOSAICrOWN has designed and developed innovative solutions to enable secure sharing and computation in multi-owner data platforms while ensuring protection of private, sensitive, or company-confidential information. MOSAICrOWN provides impact by: i) supporting the enforcement of privacy legislations and regulations, hence increasing citizens trust and control; ii) increasing the possibility of data sharing and processing in full respect of privacy and confidentiality, hence providing for better value-creation from personal and proprietary data; iii) enabling new data market opportunities, supporting European digital economy; iv) demonstrating the advantage of developed solutions through appropriate use cases.

In addition to the impact given by the direct exploitation and deployment of MOSAICrOWN solutions by industrial partners, MOSAICrOWN has also achieved impact through several dissemination, communication, and exploitation-enabling activities. Also, MOSAICrOWN participates in the Big Data Value PPP partnership and has contributed actively to its initiatives.

The tools and techniques produced by MOSAICrOWN have contributed to the realization of digital data markets aligned with the democratic principles of the European society, facilitating the realization of the fundamental right of citizens to have guarantees on data protection.
MOSAICrOWN's enriched data market

Last update: 22 July 2022  
Record number: 648494  
Permalink: https://cordis.europa.eu/project/id/825333/reporting  

© European Union, 2022