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

Reporting

Project Information

<table>
<thead>
<tr>
<th>MOSAICrOWN</th>
<th>Funded under</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grant agreement ID: 825333</td>
<td>H2020-EU.2.1.1.</td>
</tr>
<tr>
<td>Project website</td>
<td>Overall budget</td>
</tr>
<tr>
<td></td>
<td>€ 3 203 750</td>
</tr>
<tr>
<td>Status</td>
<td>EU contribution</td>
</tr>
<tr>
<td>Ongoing project</td>
<td>€ 3 203 750</td>
</tr>
</tbody>
</table>

Start date 1 January 2019

End date 31 December 2021

Coordinated by

UNIVERSITA DEGLI STUDI DI MILANO

Italy

Periodic Reporting for period 1 - MOSAICrOWN (Multi-Owner data Sharing for Analytics and Integration respecting Confidentiality and Owner control)

Reporting period: 2019-01-01 to 2020-06-30

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 will provide 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 will pursue 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 meets 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. Requirements from the three representative use cases from industrial partners in MOSAICrOWN have been gathered and analyzed. The work (focus of WP2 "Requirements") was performed in strong cooperation among all the partners to discuss the collected requirements, identifying commonalities and peculiarities of each use case, and to analyze and organize the requirements. The approach was iterative with refinement phases to produce the
final list of requirements to be considered by MOSAICrOWN. The requirements have provided input to the research activity and are also being considered for ensuring alignment with research and deployment of technical solutions to the use cases.

Objective 2 - Data governance framework. The work on this direction (focus of WP3 - "Data governance framework") has addressed the definition of the overall architecture of the data governance framework and of the policies regulating its behavior. The work has investigated possible architectural solutions for the governance framework and has produced a first version of the 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 work has also investigated possible directions for the realization of the policy enforcement engine.

Objective 3 - Data wrapping. The work on this direction (focus of WP4 - "Data wrapping") has addressed the problem of protecting data stored in the data market, also with consideration of distributed settings. The proposed solution combines the All-Or-Nothing-Transform for strong resource protection and resource fragmentation with decentralized allocation. The work has also investigated solutions enabling data owners to contribute to the digital data market while maintaining control over their data, considering in particular scenarios of collaborative query executions and data trading.

Objective 4 - Data sanitization. The work on this direction (focus of WP5 - "Data sanitization") has investigated privacy and utility metrics relevant for digital data market scenarios, and developed sanitization solutions considering both syntactic and semantic approaches. The first version of tools released provide for an efficient k-anonymity evaluation leveraging a distributed architecture and for a differentially-private sanitization able to resist membership inference attacks in machine learning applications.

Objective 5 - Effective exploitation. Industrial partners have designed and pursued exploitation presenting MOSAICrOWN and some preliminary 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 is working on additional anonymization techniques (aiming to extend HANA’s capabilities), and privacy interpretations for machine learning in business applications; EISI is integrating 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 continues to refine its data security practices and is investigating the integration of MOSAICrOWN techniques into its day-to-day operations and offerings.

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 develops 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 also achieves impact through several dissemination, communication, and exploitation-enabling activities. Also, MOSAICrOWN participates in the Big Data Value PPP partnership and contributes actively to its initiatives.

The tools and techniques produced by MOSAICrOWN contribute 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.

_Last update:_ 27 November 2020  
_Record number:_ 648494