

Input Privacy-preserving Techniques

Type of Activity			
<input type="checkbox"/>	New project	<input type="checkbox"/>	New activity
<input checked="" type="checkbox"/>	Extension of existing project	<input type="checkbox"/>	Extension of existing activity
The project is an extension of the 2020 Input Privacy-Preserving Techniques project sponsored by the HLG-MOS.			
For an overview of the progress of the current IPP project, refer to the meeting documents: https://statswiki.unece.org/display/IPPP/Meeting+Documents			
Purpose			
<p>Statistical organizations are more and more investing on becoming part of a data ecosystem where they acquire and integrate data from multiple sources and provide richer statistical products.</p> <p>In this scenario, the issue of privacy preservation is particularly relevant: the more sources are acquired and integrated, the higher are the risks of disclosing information violating individual privacy rights. Hence, from a legislative perspective there are indications to take privacy into account throughout the whole data treatment process, through the 'privacy by design' concept.</p> <p>National Statistical Organizations (NSOs) are used to apply techniques for enforcing privacy by design on the output side, i.e. when publishing aggregated statistical data for dissemination purposes and when sharing microdata for research purposes with statistical disclosure control (SDC) and other output privacy-preserving techniques.</p> <p>However, NSOs have still to invest on dealing with privacy protection on the input side, in a complementary but distinct way with respect to output privacy preservation investments¹.</p> <p>Different classes of techniques can be used to deal with input privacy². Among them Secure Multiparty Computation (SMC) and Homomorphic Encryption (HE) play a relevant role.</p> <p>These methods are particularly suitable for use in a non-trusted environments such as access to private data, interconnectivity of highly sensitive data for the purpose of scientific research, data analytics in Cloud and AI. The goal of this project is to investigate statistical use cases that require protection on the input side, assess and determine applicability of selected classes of techniques for main scenarios, identify opportunities for sharing across statistical community and create community of practice across statistical organizations and external partners (academia, private sector).</p>			
Description of the activity			
<p>After a late start, a number of use cases in the field of input Privacy-preserving techniques were inventoried and documented in 2020. In 2021, the focus will be on gathering experiences from generalized use cases. External technical audiences will be called for help. The following work packages have been identified:</p>			

¹ F. Ricciato, A. Bujnowska, A. Wirthmann, M. Hahn, E. Barredo-Capelot, A reflection on privacy and data confidentiality in Official Statistics, ISI 2019.

² UN Handbook on Privacy-Preserving Computation Techniques, <http://publications.officialstatistics.org/handbooks/privacy-preserving-techniques-handbook/UN%20Handbook%20for%20Privacy-Preserving%20Techniques.pdf>

- **WP1: To document, generalize and prioritize the use-cases:** New participants will bring new use-cases that will be added to the inventory and possibly generalized. A subset of most prominent use-cases will be identified. Reference scenarios that are relevant for different types of source data and/or different NSOs will be drafted and prioritized for follow-up action in WP2.
- **WP2: Setup and test the use-cases:** Selected types of use-cases will be tested against different techniques. For use-cases for which a broader spectrum of technological solutions is available, we will launch a wider call for proposed solutions from external technical audiences. Use-case testing and comparison of different solutions based on benchmark conditions (e.g. with synthetic test data) will be conducted. Technology aspect will be considered as well as organizational, business and legal aspects. Due to the complexity of this work package, it might be structured into parallel sub-streams.
- **WP3: Lessons learned:** The experiences from WP1 and WP2, will be used to make clear recommendations on which techniques are most suitable in which cases and where such techniques are not applicable. Although the focus will be on technical and methodological aspects of implementation in official statistics, lessons learned and recommendations with respect to organizational, business, and legal aspects will be documented as well.