

January-February 2020

Machine Learning

Progress

- The **Project participants** and **Contributors and followers** list have been updated. Some have moved from one list to other. Some people have been dropped from the Contributors and followers list. Several people and organisations have joined the project. The project currently has 38 participants from 18 organisations and 14 countries, and 42 followers and collaborators from 24 organisations and 17 countries.
- Working meetings, presentations and other collaborations continue to take place. Monthly progress and update meetings are attended by 25 to 30 team members and most organisations involved.
- Project outputs, delivery dates and tasks have been set. The outputs are: pilot study reports; summary reports on the application of ML in three areas (Coding and classification; Edit and imputation; Imagery); summary report on the integration of ML in production processes (including best practices, quality framework); material to deliver project outputs through a live event, e.g. workshop; and, a project wrap-up and recommendations on any next steps.
- Two new pilot studies have joined the project. One Coding and classification from the IMF has already started to obtain the knowledge, experience and code that they have developed in recent months. The other project is from Istat on editing. Other discussions are taking place on adding pilot studies.
- The scope and activities on work package 3 on organisational aspects is being discussed and finalized.
- The pilot studies continue to progress. Participants have until the end of February to deliver a first version of their reports.
- The UNECE Secretariat is leading discussions with the UN Global Platform as one of the means to share the outputs of the project. A small test has confirmed the project's ability to share some ML code on the platform. A more substantive test is planned. The project will use a combination of the UNECE wiki, Github and the UN Global Platform to share its outputs.
- A sprint to be held in Warsaw on April 1-3 is being planned. Participants have until the end of February to indicate their intention to attend. Participation is strongly encouraged as the sprint will be crucial to complete the first outputs of the project and put the others on the right path. So far, 10 countries have indicated that they will attend; others will participate via Webex. Participation is strongly encouraged.
- Some organisations have been contacted to obtain information on ML applications that have been demonstrated or have been integrated to production. This information will enrich the analysis on the added value of ML and provide valuable insights on their different stages towards integration.
- A session including 4 papers from the ML project was submitted for the BigSurv20 conference in Utrecht from November 4 to 6.

Next Steps

- Finalize the planning of 2020 activities, including identification of new pilot studies, setting the scope and activities of WP3, deciding on the platform to share the outputs of the project.
- Obtain the first version of pilot study reports by the end of February.
- Submit a session at the BigSurv2020 conference in Utrecht from November 4-6 to showcase the main outputs of the ML project.
- Set the agenda of the April sprint.

Risks and Issues

Issue	Mitigation

Input Privacy-preserving Techniques

Progress

blocked URL

Limited progress has been made. Countries have been requested to provide project managers but so far none have been identified. Statistics Netherlands will look into providing one part time project leader or substantive specialist to co-lead the project.


Lack of resources at UNECE has further hampered the start of the project.




Next Steps

Risks and Issues

Issue	Mitigation

News from the Groups		
<p>Blue-skies Thinking</p> <p>? Unknown Attachment</p>	Identifying Topics / Opportunities	IN PROGRESS
	Follow-up selected topics	IN PROGRESS
	Other	
<p>Capabilities and Communication</p> 	Culture Change and Internal Communications Strategy	<p>IN PROGRESS</p> <p>Sprint session on Modernisation Group on Developing Organisational Capability and Communication was organised on 4th February 2020. In the past 2 years we had two phases of the Strategic Communication Project. It was decided to add these issues back to our modernisation group. The sprint session was focused on identifying areas of work related to communication and defining activities for this year. It was agreed to establish three joint Task Teams on: internal communication and culture change, social media strategy and other follow up on the Strategic Communication Framework, training on staff in communication. Now, we are waiting for volunteers for various TFs.</p>
	Competencies Training and Development	IN PROGRESS
	Future of work in the context of Modernisation of the workplace	IN PROGRESS
	Ethical leadership	IN PROGRESS

	Follow-up Strategic Communication Framework	IN PROGRESS	
	Training of staff in communication	IN PROGRESS	
	HR MT Workshop	IN PROGRESS	
	Other		
<div>Supporting Standards</div> 	Linking GSBPM and GSIM	IN PROGRESS	<p>The Task team started working on other GSBPM phases, namely phase 2 - Design, phase 4 - Collect and phase 6 Analyse. In total, 12 sub-processes are currently under development by one or two countries. In 2019, the team completed phase 5 - Process of the GSBPM.</p> <p>The task team discussed how to speed up work on sub-processes without losing the benefit of having two (or more) countries working on the same sub-process. The team decided to have more interaction and cooperation among countries that work on the same sub-process so that they could work together more intensively and effectively and save time and resources. Indeed, to fill in the specification and implementation level of a single sub-process implies a lot of work for countries.</p> <p>The first mappings produced in January were discussed during the February meeting.</p> <p>The task team is also revising the document summarising the theoretical approach developed for the mapping in 2019 and the examples developed by the task team for the phase 5 - Process of the GSBPM.</p>
	Core Ontology for Official Statistics	IN PROGRESS	
	GSBP M for Geospatial data	IN PROGRESS	<p>During the first two meetings, the team investigated various examples of geospatial information within statistical organisations and agreed to first start working on how to use geospatial information to produce geospatially-enabled statistics using GSBPM.</p> <p>The team will use GEOSTAT2 project (2015-2017; Eurostat) output as a starting point, review its final report (which contains geospatial-related activities from GSBPM phase 1 to phase 7) in consideration of recent developments (e.g. Global Statistical and Geospatial Framework (GSGF) from UN-GGIM).</p> <p>One of primary objectives of the first two meetings was to identify other on-going initiatives in this field and coordinate with them to avoid duplication of efforts.</p> <p>Kick-off meeting was attended by Eurostat's GEOSTAT4 project team (Statistics Finland and Eurostat) and UN-GGIM's Expert Group on Integration of Statistical and Geospatial Information (Germany).</p> <p>It was decided that the task team will work on integration using GSBPM as described earlier and once the work is done, GEOSTAT4 project team will review it so that work of the two teams doesn't overlap.</p>

	Metadat a Glossary	IN PROGRESS	<p>The Task team has concluded its work.</p> <p>During the February meeting, the task team completed the review of GSIM terms. The task team will meet again in March for final fine tuning of the work done.</p> <p>The Metadata Glossary team developed definitions and explanatory notes for terms in GAMSO and GSBPM and revised all the terms of the GSIM Glossary to verify their compliance to formal rules (for terms definitions), the clarity and the content of the explanatory texts.</p>
	Other		<p>As previous two years, Supporting Standards Group is organising together with Sharing Tools Group the ModernStats World Workshop 2020 from 24 to 26 June. The Statistical Office of the Republic of Serbia offered to host the Workshop. The invitation letter has been sent out. The organising committee was set up and have the first meeting end of February.</p> <p>The other activity foreseen for 2020 is the light update of GSIM. The task team will start working after Metadata Glossary team concludes its work. Members who indicated interest will be contacted before April for preliminary scoping.</p>
<div>Sharing Tools</div> <div>? Unknown Attachment</div>	Digitizing/editing CS PA document & Communication restated CS PA	IN PROGRESS	Pending
	Adding Services to Catalogue	IN PROGRESS	Services from the ESSnet I3S project will be added. Algorithms coming from the Machine Learning project can also be included as long as sufficient documentation in line with the CSPA requirements are provided.
	Integrating Innovations in Architecture and technology	IN PROGRESS	Results from the ESSnet sprint in Toulouse will be checked and where relevant findings will be included in the CSPA concept
	Other		The group is still lacking a chair (or co-chairs). UNECE is lacking resources to provide the necessary support to the group

1. [HLG-MOS \(1 March, New York\)](#)
2. [Machine Learning Sprint \(, Warsaw, Poland\)](#)
3. [ModernStats World Workshop, 24-26 June 2020, Belgrade, Serbia](#)
4. [Workshop on Statistical Data Dissemination and Communication, 7-9 September 2020, Lisbon, Portugal](#)
5. [Workshop on Human Resources Management and Training, 9-11 September 2020 Lisbon, Portugal](#)
6. [Machine Learning Sprint \(tbd, Germany\)](#)
7. [Workshop on Statistical Data Collection, October 2020 \(tbc\)](#)

8. [HLG-MOS workshop on the Modernisation of Official Statistics, November 2020 \(tbc\)](#)