

WP1 - Pilot Studies

- Under the HLG-MOS ML Project Work Package 1, a total of 19 pilot studies were conducted with three broad themes: coding and classification, edit and imputation and imagery analysis and 2 relevant studies
- Work package 1 report provides executive summary of all three application areas.
- Theme report provides overview of context, methods, practices and lessons learned from pilot studies under each theme.
- The presentations provided in the first table were delivered at project sprints. Talks delivered at organized events, such as conferences or workshops, are provided in the second table.
- Pilot study paper contains details about each study, please see [Studies and Codes](#) page for information about programming language and codes.
- In addition to the pilot studies, other developments or uses of ML were presented to project members. They are shared to further highlight the interest in advancing the use of ML, please see [Other applications of Machine Learning](#).

WP1	Pilot Study Theme	Pilot Study Paper	Presentation
Work Package (WP) 1 - Pilot Studies Executive Summary	Coding and Classification Theme Report	Mexico - Occupation and Economic activity coding using natural language processing	Presentation (November 2020)
			Presentation (April 2020)
		Canada - Industry and Occupation Coding	Presentation (April 2020)
		Belgium Flanders - Sentiment Analysis of twitter data	Presentation (April 2020)
		Serbia - Coding textually described data on economic activity collected from Labour Force Survey	Presentation (April 2020)
		USA BLS - Coding Workplace Injury and Illness	Presentation (April 2020)
		Poland - Production description to ECOICOP	
		IMF - Automated Coding using the IMF's Catalog of Time Series	Presentation (April 2020)
		Iceland - Automatic coding of occupation and industry in social statistical surveys	Presentation (April 2020)
		Norway - Standard Industrial Code Classification by Using Machine Learning	Presentation (April 2020)
	Edit and Imputation Theme Report	Italy - Imputation of the variable "Attained Level of Education" in Base Register of Individuals	Presentation (November 2020)
			Presentation (April 2020)
		Poland - Imputation in the sample survey on participation of Polish residents in trips	Presentation (April 2020)
		Germany - Machine learning for imputation	Presentation (April 2020)
		Belgium VITO - Early estimates of energy balance statistics using machine learning	
		Canada - Time Series Models for Early Estimates of Energy Balances	Presentation (October 2020)
		UK - Editing of Living Cost and Food Survey Income data	Presentation (April 2020)
		Italy - Editing in the Italian Register of the Public Administration	Presentation (April 2020)
		Italy - Machine Learning for Data Editing Cleaning in NSI : Some ideas and hints	Presentation (November 2020)
	Imagery Analysis Theme Report	Australia - Address Register Automated Image Recognition (AIR) model	Presentation (November 2020)
		Netherlands - Learning statistical information from images: a proof of concept	Presentation (April 2020)
		Switzerland - Arealstatistik Deep Learning (ADELE) ADELE page at the Swiss Federal Statistics Office	Presentation (November 2019)
		Mexico - Use of Landsat satellite data for the mapping of urban areas in non-census years	Presentation (April 2020)

		UNECE - Generic Pipeline for Production of Official Statistics Using Satellite Data and Machine Learning	Presentation (November 2020) Presentation (April 2020)
--	--	--	---

The following talks were delivered by project members are organized events, such as conferences and workshops.

WP	Topic	Title	Event	
WP1	Imagery	Mexico - Integrating EO with Official Statistics using Machine Learning in Mexico	GEO WEEK 2019 - Session on EO for Official Statistics	Presentation Video (starts at 7m45s)
WP1	Edit & Imputation	Germany - The UNECE HLG-MOS Machine Learning Project: A report of the Editing & Imputation Group	UNECE Statistical Data Editing Virtual Workshop 2020	Paper Presentation
WP1	Edit & Imputation	Italy - ML to identify patterns behind errors in STS statistics	UNECE Statistical Data Editing Virtual Workshop 2020	Paper Presentation
WP1	Edit & Imputation	UK - Editing of Social Survey Data	UNECE Statistical Data Editing Virtual Workshop 2020	Paper Presentation
WP1	Edit & Imputation	Italy - An imputation procedure for the Italian attained level of education in the register of individuals based on administrative and survey data	UNECE Statistical Data Editing Virtual Workshop 2020	Paper Presentation
WP1	Edit & Imputation	Italy - The imputation of the "Attained Level of Education" in the base register of individuals: an experimentation using Machine Learning techniques	UNECE Statistical Data Editing Virtual Workshop 2020	Poster Presentation
WP3	Integration	Poland & USA - HLG-MOS Machine Learning Project: Sharing ML Techniques and Algorithms, how to tackle large data sets and build international capability	6th International Conference on Big Data for official Statistics	Presentation Video (starts at 8h47m)
WP1	Edit & Imputation	UK - Editing of Social Survey Data with Machine Learning - A journey from PoC to Implementation	World Statistics Day 2020	Paper
WP2	Quality	Collaboration - The integration of machine learning into official statistics	BigSurv20 - Big Data Meets Survey Science	Video
WP1	Edit & Imputation	Belgium and Collaboration - Case studies on machine learning for editing and imputation	BigSurv20 - Big Data Meets Survey Science	Video
WP1	Coding & Classification	Belgium - Algorithmic choices for sentiment coding of Flemish tweets	BigSurv20 - Big Data Meets Survey Science	Video
WP1	Imagery	Australia - ABS use of machine learning to classifying addresses use on the Address Register	BigSurv20 - Big Data Meets Survey Science	Video