

Business Case for Accelerating the Implementation of ML-based Solution in Data Editing

This business case was prepared by the Applying Data Science and Modern Methods Group and is submitted to the HLG-MOS for their approval.

Type of Activity			
<input checked="" type="checkbox"/>	New activity	<input type="checkbox"/>	Extension of existing activity
Proposed Modernisation Group(s) for Activity			
<input checked="" type="checkbox"/>	Applying Data Science and Modern Methods	<input type="checkbox"/>	Blue Skies Thinking
<input type="checkbox"/>	Capabilities and Communication	<input type="checkbox"/>	Supporting Standards
<input type="checkbox"/>	<i>Other:</i>		
Purpose			
<p>Data cleansing and data editing is a field that has tested and implemented Machine Learning (ML) approaches. ML has a good potential for efficiency gains in complementing or replacing traditional methods, as well as for improving quality in ways that may be difficult to achieve with traditional methods. However, the migration from experimentation to production remains difficult for many National Statistical Offices (NSOs), especially with regard to gaining acceptance by statisticians and data users.</p> <p>It is important to explain the new methods of data editing to stakeholders within and beyond NSOs and prove that it is a good value for money to get their buy-in. It is also important to address fears and misconceptions around ML and AI, as well as other organisational barriers to implementing ML methods, including capability and technology gaps.</p> <p>This business case proposes to develop a generic guidance on key organizational issues when implementing ML-based solution in data editing in order to accelerate the productionisation process. Although the focus is on data editing, it is likely that the guidelines produced would be helpful for implementing ML projects in other parts of the statistical business process.</p>			
Description of the activity and deliverable(s)			
<p>Use cases will be collected to review how project managers attempted to manage the obstacles that have been stopping NSOs from applying ML-based solution in data editing throughout the journey from experiment to production. One of the key focuses is about multi-level engagement (e.g., the dynamics between senior management, research team, business area, IT team, etc.). The problems and stakeholders involved will be identified and the actions taken to resolve the issues will be summarised as lessons learned and best practices.</p> <p>To help contextualise the use cases, the pros and cons of applying ML-based solution in data editing will be highlighted, in particular to explainability, transparency and cost-effectiveness.</p> <p>Although the focus is on data editing, where it may be helpful, we will consider drawing on successful implementations of ML in other fields such as imputation or classification where agencies may have had early gains.</p>			

A guidance will be developed based on the use cases to offer recommendations to facilitate the implementation of ML-based solutions in data editing.

Alternatives considered

NSOs may exchange their experiences on a bilateral basis. NSOs may also engage with one another through the ONS-UNECE ML 2022 program and any successors.

How does it relate to the HLG-MOS vision and other activities under the Group or HLG-MOS?

This business case aligns with the HLG-MOS vision and values, in particular to accelerating the development of innovative solutions and openly discussing challenges and opportunities.

Proposed start and end dates

Start: January 2023

End: December 2023