

---

## **Moving towards the standardized process of automatic statistical data editing using machine learning techniques**

Ieva Burakauskaitė (Statistics Lithuania State Data Agency, Lithuania)

[ieva.burakauskaite@stat.gov.lt](mailto:ieva.burakauskaite@stat.gov.lt)

### *Abstract*

As it is well known, statistical data editing is a significant but time-consuming process during the production of official statistics at National Statistical Institutes. In order to increase the efficiency of the statistical data editing, the Generic Statistical Data Editing Model (GSDEM) offers some valuable insights on various steps during the latter process such as the detection of the most influential errors using selective editing and the error treatment with either interactive or preferably automatic editing. Although the selective editing process step had already been developed and implemented at State Data Agency (Statistics Lithuania), the error treatment automatization part still needed to be refined. However, the migration of statistical surveys into the uniform cloud-based software platform has motivated Statistics Lithuania to reevaluate the current statistical data editing process. Such a platform with all available administrative and statistical data in one place offers numerous opportunities for the standardization and integration of the process. The presentation will focus on the experience moving towards the standardized automatic statistical data editing within the integrated platform. The current selective editing process step will be presented and the results of a few machine learning techniques for automatic editing will be compared as a case study example.