

AIML4OS WP 11: USING ML FOR ESTIMATING FIRM-LEVEL SUPPLY CHAIN NETWORKS

Presented By

Gert Buiten – Statistics Netherlands

Supply chain networks data matter!!





KEY DRIVERS AND OBJECTIVES

Key Drivers

- Supply chains are ever more important, but complete firm-level datasets are only available for some countries
- Some EU Member States have rich data, others not due to differences in national VAT-laws
- So: use data from MS's with rich data to help others create 'fit for purpose' datasets



BASIC APPROACH

- Build a firm-level network dataset on rich Portuguese administrative data and use that for deriving training, test and validation sets
- Develop and train Machine Learning models that can estimate buyer/supplier relations from data that is available at all NSI's (building on academic work from Oxford University/INET)
- Apply the reconstruction models in several countries to test their performance and general applicability – and improve the models and software
- Develop methods to add weights to the estimated links
- Make the resulting software available to other MS's in an 'actionable way'



POINTS OF ATTENTION IN MODELLING

- Models should at least include NACE-code, geographical distance and firm size
- Feature engineering in such a way that the models can be applied in all EU countries (e.g. normalization of geographical distances and firm sizes)
- Find the best way to deal with the fact that a buyer/supplier link is a 'rare event' (<1% of potential links actually exists)
- Create robust coding pipelines for the software
- Develop adequate quality indicators



THE WP11 TEAM

 National Statistical Institutes: Portugal, Poland, Netherlands, Ireland

 Universities: Free University Brussels and Oxford University (INET)

Eurostat

6



SCHEDULE

Training and test sets
Final Models
Report on 'binary networks'
Report on 'weighted networks'
Mid 2027
Publication/dissemination
Karly 2028

