BSTN activity proposals for 2021

UNECE HLG-MOS Blue Skies Thinking Core Team, November 2020

Introduction

During the summer of 2020, the BSTN core team has organized a number of dedicated sessions where members pitched their ideas for joint activities that might be undertaken during 2021, given sufficient interest. These pitches have led to the activity proposals below that we would like to discuss in the HLG-MOS workshop 2020. Being proposals, they are open for discussion and amendment. Depending on appetite we can decide how to proceed and organize follow-up.

Some activity proposals have logical connections with other HLG-MOS related work. In principle, the BSTN core team is willing to take care of launching and overseeing the activities proposed; but if another group is interested in adopting an activity that could be an option as well.

All proposals are presented in a brief and concise way that hopefully provides enough insight to judge what the idea is about. The name of the person who pitched the idea is included; he/she can provide more content and context if desired.

Key questions to be addressed during the HLH-MOS workshop breakout session include:

- What do you think of the activity proposal as such?
- Is the proposal clear? Does it generate any further ideas/thoughts?
- Would you (and/or colleagues) be interested in participating/contributing to follow-up?
- How should we organize follow-up (if we go ahead)?

Given their different nature, these questions apply to each proposal separately.

This is the first time we present BSTN activity proposals in this way. We therefore welcome feedback on the general approach and presentation.

Very impressive my dear colleague- but does it also work in theory?

The power of network data: feasibility studies (Daniel Elazar, ABS)

Network data (e.g. from administrative registers, mobile phone calls, financial transactions, internet traffic or social media) provides a wealth of new opportunities for addressing difficult policy problems in economic performance, labour market supply and demand, energy supply and demand, transport and infrastructure and social resilience. A few examples are job flows and it's impact on regional economies, population mobility, freight transport and domestic business supply chains. However for many NSOs, obtaining unit level network data from data providers is quite difficult due to a number of barriers such as privacy concerns, commercial interests, policy and legislative restrictions, social license, data transfer and ingestion limits and technical capability. A small but growing statistical literature exists in partially observed networks and inferring network connectivity, which could form the basis for an empirical evaluation study. The immediate goal of this BSTN activity is to determine the methodological feasibility of deriving sufficiently valid inferences about networks from data aggregates in a manner that addresses most of these concerns. If methodological feasibility can be demonstrated, next steps will involve choosing suitable business applications that are of interest to all participants and consulting with network data providers and analysts on the feasibility of taking this forward.

COVID-19 Hotspot Joint Biosecurity Centre Platform, a secure environment for big data and data science collaboration (Eric Deeben, ONS)

The Joint Biosecurity Centre (JBC) was created in helping to break the chains of COVID-19 transmission through expert analysis and insight, enabling evidence-based decisions at a national and local level. ONS Data Science Campus and Digital Service Technology build a flexible analytical single platform for all data and analysts to work in one place, assured for data security but flexible enough to adapt to changing needs, built in Google cloud for scalability. Cloud native tools for data management and modelling, outputs are managed and controlled, access is authenticated. Data processing in Big Query, Python, R or a combination of them all; use the latest, most powerful libraries and modules (including geography and NLP).

User Research for Official Statistics (Eric Anvar, OECD)

The idea follows from questions raised in the StatsBot activity group during 2020, like: *Which are the users' problems we are trying to solve? What is our level of understanding of users and users' needs? Are we at edge of user research techniques available and used on the market?* Because statistical organisations often serve similar users, we believe there is an opportunity to try to answer such questions together, adopting advanced user research techniques – including AI based ones – like *Lean Start-up* and evidence-based product design approaches. We could start by assessing our maturity vs. modern user research techniques. Further ideas for joint activities include developing a 'statistical personas' database, a research study on data journalists, and a toolkit to improve visibility in search engines. Note that this activity proposal differs from previous work in e.g. the strategic communications project in that it aims to focus on concrete user-oriented research in an international context rather than on developing strategy.

Rapid survey system (Branko Josipović, SORS)

The COVID19 outbreak imposed many difficulties in regular data collection. These difficulties were very visible in the domain of business statistics. During this unpredictable time, it was very hard to collect all necessary information regarding the situation in the economy depending only on regular surveys or administrative sources. In addition, decision-makers articulated additional requests for statistical information to support economic forecasting.

The Serbian statistical system had to be adjusted to cover the new situation. On the one hand we needed to maintain different communication channels with reporting units as a key source of data. At the same time, official statistics had to be able to quickly respond to unforeseen user demand. The best way for us to do that was to introduce new and rapid surveys, which helped us to deal with the situation i8n the best possible way. This type of survey, nevertheless, must be carefully designed to achieve time efficiency, data quality and not to overburden reporting units at the same time. Would it be worthwhile to share lessons learned and best practices on similar experiences?

From experimentation to implementation in official statistics (Kate Burnett-Isaacs, StatsCan)

Many NSOs are forging into newfound territory of data science and analytics. These new modern methods and tools have expanded the possibilities for producing more timely and granular statistics and reaching broader audiences and uses of statistical information. While researchers in these organizations have made significant grounds on applying these methods to pilot projects, moving these new technologies and methodologies into production of official statistics has been met with its own set of challenges. Hurdles to move from experimentation to implementation in a production environment are not limited to technological capacity or knowhow, but also require a transformation in thinking and culture within the agency's policies and workforce. In order to capitalize on the benefits of new methods to producing statistics, statistical offices around the globe have taken different approaches to integrate data science and analytics into their day to day work. This activity will explore the environments of participating NSOs, the methods and technology they are adopting into their official statistics and the means they have used to make data science and analytics a part of the culture within their agency. The goal is to share experiences, lessons learned and open a discussion on ways to collaborate internationally to make it easier to adopt new methods and technologies within our own organizations in the future.

Microdata for understanding falling response rates (Gary Dunnet, StatsNZ)

NSOs across the world are facing the challenge of falling response rates, while demand for data is growing, and while admin data offers some reprise face-to-face interviewing will remain for a long-time to come; so it seems that sharing our experiences is critical if we are to deliver an effective and efficient service. The paper "Survey Nonresponse Trends and Fieldwork Effort in the 21st Century: Results of an International Study across Countries and Surveys (Journal of Official Statistics, Vol. 36, No. 3, 2020, pp. 469–487, Annemieke Luiten, Joop Hox, and Edith de Leeuw)" goes some-way in sharing experiences, but they note that "*A striking result is that none of the fieldwork effort variables is related to overall response, while most of the general design variables are related to response trends – but micro-level data availability was an issue*", and so my pitch is that we (HLG-MOS) commission work to take this research further, in consultation with the original authors and support them in their quest to get the data they need to draw this work to a conclusion.