Blue Sky Thinking Network –
Looking Back & Ahead

By UNECE Modernstats, November 2018
What were the BSTN goals?

• Identify and evaluate new opportunities
• Generate new ideas
• Develop concrete proposals
• Coordinate internationally
• Create synergy with other platforms
Looking Back at 2018

Success
• BSTN 2018 has resulted in
  - Machine Learning Position Paper
  - ML Project Proposal
  - Plan to work on similar IT position paper(s) in 2019

Learning
• “Think Tank” role of BSTN not (yet) fully realised
  - *Do we need a change to improve this?*
This paper is written for managers and policy makers to inform them about the possibilities to use machine learning (ML) in the production of official statistics and to demystify ML for those official statisticians unfamiliar with it.
ML position paper (2/4)

- Contents (14 pages, we aimed for 10…)
  1) Introduction
  2) Conclusions and recommendations
     (Next slides)
  3) Background
     (What is it, why relevant?)
  4) ML in primary data
     (GSBPM based)
  5) ML in secondary data
     (Big data, register data, multisource)
  6) Quality considerations
     (Traditional indicators vs. ML-specific issues)
  7) References
• Conclusions

✓ Processing of secondary data is steering NSOs towards use of ML
✓ But use can also extend to more traditional domains

<table>
<thead>
<tr>
<th></th>
<th>Family of ML techniques</th>
<th>GSBPM phase</th>
</tr>
</thead>
<tbody>
<tr>
<td>Record linkage</td>
<td>Clustering</td>
<td>2.4, 5.1</td>
</tr>
<tr>
<td>Coding</td>
<td>Classification</td>
<td>2.4, 4.3, 5.2</td>
</tr>
<tr>
<td>Outlier detection</td>
<td>Clustering</td>
<td>2.4, 4.3, 5.1, 6.2</td>
</tr>
<tr>
<td>Stratification</td>
<td>Classification</td>
<td>4.1, 4.3, 5.4, 5.6</td>
</tr>
<tr>
<td>Estimation</td>
<td>Regression/classification</td>
<td>4.3</td>
</tr>
<tr>
<td>Imputation</td>
<td>Regression/classification</td>
<td>5.4</td>
</tr>
<tr>
<td>Calibration</td>
<td>Regression/classification</td>
<td>5.6</td>
</tr>
<tr>
<td>Disclosure control</td>
<td>Regression/classification</td>
<td>6.4</td>
</tr>
</tbody>
</table>
Recommendations

- Develop quality framework
  - New, or integrated in existing one?
  - Possibly also guidelines for ML use
- Create inventory of ML projects
  - To exchange both methods and applications
  - Not one-off but living document
  - Including previous work (Chu&Poirier, Destatis)
- Set up interdisciplinary teams for ML work
  - Modelers
  - Programmers
  - Subject matter specialists
Next year- new IT trends

- Group on new IT trends like
  - Secure multiparty computing
  - Edge computing
  - Blockchain techniques
  - Cloud services

- For each topic short position paper (aim 3 pages, ready in 3 months)
  - What is it
  - What does it mean for OS
  - What is already being done
  - What future work is needed

- Start with physical meeting?

- Core team: Canada, Italy, Netherlands, UK
  - Others welcome!
What more to promote BSTN?

- Invest in building a real network and core group
- Combine top-down and bottom-up approaches
  - Monitor relevant ideas from high-level meetings
  - See what’s happening elsewhere (i
  - Collect suggestions from HLG-MOS workshops
  - Liaise with HLG-MOS Modernisation Committees
  - Get the people within NSO’s innovation labs together
- Arrange meeting(s) where hands-on blue sky thinkers:
  - Share BST happening within their organizations
  - Share, discuss and prioritize any BST ideas that might be promising
  - Agree on (limited!) next steps to realize ideas
  - Translate ideas into proposals
- Liaise with other communities
  - Geo, academic world, governments, NGOs, private sector, …
- Allocate resources to develop and work on BST ideas
- Secretarial support from UNECE
What you can do:

- Join the Network
- Bring your ideas
- Contribute to ideas
- See what is happening

Contact:
- Innovation Manager: Barteld Braaksma (bbka@cbs.nl)

Wiki platform: https://statswiki.unece.org/display/BST
Blue Sky Thinking Network – Looking Back & Next Steps

By UNECE Modernstats