

## SFI Public Service Fellowship 2023

<b>1. Name of Governmental Department or Agency</b>
Central Statistics Office
<b>2. Title of the Project</b>
<b>CSO4</b> Nowcasting CO <sub>2</sub> Emission and Trade Routes in Real-Time using Big Data
<b>3. Description of the Project</b>
<p>The COVID-19 pandemic highlighted the demand for timelier data to inform policy and decision-making. Using big data in official statistics can generate those desired real time information products. Big data is a potential source of data for the CSO to compile official statistics with better timeliness, granularity, coverage, and accessibility compared to more traditional data sources.</p> <p>This project focuses on Automatic Identification System (AIS), an innovative technology that offers the opportunity to derive nearly real time shipping data to generate faster transport indicators. AIS is the international system for tracking ship movements which was originally developed by the International Maritime Organization (IMO). Ships of a certain size and type are required to carry an AIS transponder which transmits their location and other key information in real time.</p> <p>Stored AIS data is an example of big data due to the extremely large amount of signal data generated every minute by ships around the world. It is a promising novel data source which offers potential for statistical analysis going beyond its original goal of maritime safety.</p> <p>Ocean shipping is the main mode of transport for international trade activities with vessels carrying around 80% of the volume of global trade. Therefore, the production of maritime indicators in official statistics using AIS is vital for the understanding of trade activities. Furthermore, Ireland's Climate Action Plan for 2030 will see an increase in the need for more timely transport indicators. To reduce the amount of CO<sub>2</sub> emitted by the maritime transport sector, the IMO set a strategy to reduce at least 50% of CO<sub>2</sub> emissions, compared to 2008 levels, by 2050. However, there is no prescriptive measure to reach the goal. In 2018, global shipping emissions represented over 1 000 million tonnes of CO<sub>2</sub>, and were responsible for nearly 3% of global emissions caused by humans.</p> <p>The aim of this project is to use AIS data to analyse trade routes and to estimate shipping emissions. Those indicators should be of high enough quality to be published as official statistics. This project expands the understanding of the role big data can play in the production of official statistics.</p>
<b>4. Project Scope</b>
<p>The AIS project has the following overall strategic goals, which the SFI Fellowship researcher will contribute to:</p> <p><b>1. Nowcasting trade routes in real-time using AIS data:</b></p>

Modern technology like AIS for vessels enables us to track trade flows in real time. Consequently, data coming from AIS data have the potential to serve as a fast and granular indicator for trade activities. The project will use official statistics from Ireland as a benchmark to evaluate if AIS data can be used to generate official real time statistics. The project will use machine learning (ML) to analyse trade routes using AIS data. The aim is to develop code that automatically publishes results on to the CSO Transport Dashboard. This analysis will generate independent insight on maritime trade in real time.

**2. Estimating shipping CO2 emissions using AIS data:**

To measure Ireland’s ship emissions, an estimation model will be developed.

The aim is to produce a replicable method to estimate emissions. It will expand on the first goal and adds the emissions of worldwide shipping for both domestic and international trade. The analysis will be directly relevant to measure progress on Ireland’s Climate Action Plan for 2030.

**3. Contribute to the understanding of the usage of big data in official statistics:**

Big data sources (including AIS) are not purposely designed for statistical analysis. Hence, for the CSO, a key question remains on how the quality of official statistics can be guaranteed when using new big data sources. Maintaining data quality; gaining and validating methodological transparency and ensuring long-term stability in the measurement and production processes are key challenges when using big data. This project adds to the understanding of which role of big data can play in the production of official statistics and if big data are of high enough quality to generate official statistics.

For this project AIS data will be supplied by the Task Team on AIS Data of the United Nations Committee of Experts on Big data and Data Science for Official Statistics (Task Team on AIS Data — UN-CEBD) and accessed through the UN Global Platform — UN-CEBD (UNGP) which holds a global repository of live and archived AIS data.

**5. Skills/Expertise Required**

The following technical and statistical skills are needed for this project:

- Expertise in data science and being able to use machine learning algorithms to analyse big data
- Understanding of how to clean, manipulate, visualise and interpret big data
- Coding experience in python are essential as the project must be run on the UNGP which operates in python
- Ability to work effectively on own initiative and as part of a team

**6. Expected Outputs of Project**

Deliverables:

As a 6-12 months full-time project, the AIS project has the following objectives:

1. Detailed python code available on GitHub on how to generate trade routes using AIS data
2. Detailed python code available on GitHub on how to estimate shipping emissions using AIS data

3. A working code that automatically updates the Transport Dashboard on CSO Website
4. A report with illustrative examples on how new big datasets may be used on wider official statistics
<b>7. Working Arrangements</b>
The researcher would ideally be based in the offices of the CSO in Cork. Flexible and remote working arrangements will be accommodated.
<b>8. Expected Timeline</b>
The project is expected to have a 6-12 months full-time timeline with the following components:
<ol style="list-style-type: none"> <li>1. Research and Development Phase</li> <li>2. Testing and Deployment Phase</li> <li>3. Documentation and Dissemination Phase</li> </ol>
<b>9. Contact Details</b>
Dr Nele van der Wielen CSO Statistician Transport Statistics