ESSnet Big Data

Specific Grant Agreement No 2 (SGA-2)

https://webgate.ec.europa.eu/fpfis/mwikis/essnetbigdata
http://www.cros-portal.eu/.........

Framework Partnership Agreement Number **11104.2015.006-2015.720**

Specific Grant Agreement Number **11104.2016.010-2016.756**

Work Package 1

Web scraping / Job vacancies

Deliverable 1.1

Strategy for ongoing engagement (SGA-2)

Version 2018-23-03

Prepared by: Nigel Swier (ONS, UK)

ESSnet co-ordinator:

Peter Struijs (CBS, Netherlands)
p.struijs@cbs.nl
telephone   : +31 45 570 7441
mobile phone   : +31 6 5248 7775

1. Introduction

During the scoping of Big Data ESSNet SGA2 WP1, an issue was raised about how the outcomes from this work package would be taken forward in the longer term. It was envisaged that a strategy for long term engagement would be needed to ensure that what was learned by WP1 during the Big Data ESSNet could continue to be developed and disseminated to other countries looking to make use of on-line job vacancy (OJV) data. SGA2 committed the work package to the following action:

- *Develop and implement a strategy for ongoing engagement and development on the use of web scraped job vacancy data for statistical purposes within the ESS. It is envisaged that this may be the form of network of interested parties that enables cooperation to continue after the end of the ESSNet.*

To a large extent, such a strategy is no longer needed. The ESS Big Data Steering Group has agreed that the second Big Data ESSNet - starting in November 2018 and lasting for two years - will include a strand of activities to move several pilots from the first ESSNet into implementation. The WP1 pilot has been identified as one of these pilots. Therefore, the strategy has effectively already been determined. Ongoing engagement and development on the use of OJV data within the ESS will be an integral part of the second Big Data ESSNet.

However, there are some other factors that may affect ESS work on OJV data over the next few years. An important factor is the role of the European Centre for the Development of Vocational Training (CEDEFOP) as a partner and likely future source of web-scraped OJV data. Another factor uncertainty around what implementation, what it means and whether NSIs will have the infrastructure to produce Big Data statistics within these timescales. Finally, there is an issue around the future leadership of the work package. Therefore, apart from fulfilling the contractual requirement to produce this deliverable, it useful to discuss how these factors will shape how the future development of this work.

2. Factors Affecting the Future Strategy

2.1 CEDEFOP:

Over the past two years there has been a recognition of the overlap between CEDEFOP's pan-European project for gathering data about on-line vacancies[1] and the WP1 Big Data ESSNet, which aims to look at the feasibility of producing new and improved statistics from web scraped data. Both projects are EU funded and both aim to improve policy-making in relation to the European labour market. There is a recognition that these projects should avoid duplicating effort and work together to develop and apply different areas of expertise. To this end, a partnership agreement has been established between CEDEFOP and Eurostat to coordinate activities and to share knowledge. For example, a joint workshop

---

[1] http://www.cedefop.europa.eu/sl/events-and-projects/events/real-time-labour-market-information-and-skill-requirements

involving the ESSNet was held in March 2018 to validate CEDEFOP's pan-European system prior to launching continuous data collection.

The decision to take the WP1 pilot work into implementation as part of the second Big Data ESSNet presents a significant opportunity to further embed cooperation between CEDEFOP, Eurostat and the wider ESS. It is expected that the first OJV data from the CEDEFOP vacancy scraping system from selected countries will start becoming available towards the end of 2018. It is hoped that data from the CEDEFOP will become available for use by NSIs within the ESS and will become an important, and possibly, the main source of OJV data for the ESS. It is also expected that NSIs will contribute by providing statistical expertise as well as other data sources to validate data from the CEDEFOP system. These data source would include Job Vacancy Survey (JVS) data and OJV data obtained from other sources.

The call for the second Big Data ESSNet is expected in May 2018. When scoping the second Big Data ESSNet, there will be an opportunity to consider how to further deepen collaboration with CEDEFOP. This could include for example, specific actions to incorporate OJV data collected by CEDEFOP into statistical production within the ESS.

One concern is whether the European focus of the CEDEFOP system could mean that the data available will not the specific requirements of NSI. For example, job vacancies would only be available for NACE and not for country specific classifications. Although this is true, it is considered that the CEDEFOP data would enable country specific classifications to be derived where this is required.

2.2 The meaning of "implementation":

At a recent face-to face meeting of WP1, some concerns were raised about the future of an implementation project based on OJV data. A major concern is that the methods for producing sufficient robust statistical outputs based on OJV data are not sufficiently mature. Another concern is that few NSIs seem to be on track to have the necessary IT infrastructure in place to be able to produce Big Data statistics in a production setting. Finally, there is concern that there is insufficient support from subject domain experts to produce official statistics from OJV data. For this reason, any implementation project needs to be framed in terms of something that is achievable.

2.3 Work package leadership

The Big Data ESSNet WP1 has been led by the UK. Following the UK Government's intention to honour the result of the June 2016 Referendum on EU membership, the UK NSI (ONS) is currently unable to bid for EU funding. Specifically, EU funding cannot be sought for projects with time scales that extend beyond the planned departure of the UK from the EU in March 2019. The current ONS position is that it would like to be involved in second Big Data ESSNet as an informal partner. However, it would probably not be feasible for ONS to continue to lead this work package.

Given that the scoping of the second Big Data ESSNet will likely occur from May through to the end of the summer, a decision of future leadership of the work package may need to be made quite soon. WP1 partners should start considering if they might be able to lead this work package.

3. Conclusion

The overall strategy for continuing the work of WP1 after the end of the first Big Data ESSNet is clear. Over the long-term this work will continue as part of the second Big Data ESSNet. In the short to medium term (i.e. between May 2018 and November 2018) the focus will be on scoping activities for the second ESSNet. The second ESSNet work package will look to move some of the work of the WP1 pilot into production, although this needs to be framed as something that can be realistically achieved. Another country will need to step forward to lead the work package and there will need to be continued dialogue with CEDEFOP to ensure that implementation plans are coherent.