

This programme
receives financial
support from the



Norwegian Ministry
of Foreign Affairs



Workshop

Can we use Big Data for Skills Anticipation and Matching?

19-20 September 2019
Room VI, R3 South, ILO



Background

In dynamic and constantly changing labour markets, identifying skills needs is an important challenge. Imbalances on the labour market, reflected by difficulties faced by businesses in sourcing the skills they need, high incidences of skills mismatch, and significant unemployment or underemployment especially among youth, are common to most countries. In order to tackle these issues, policy-makers, employers, workers, providers of education and training and students all need timely and accurate information about demand for skills on the labour market. Traditionally, policy-makers have used information both from official labour force surveys, and from other surveys to provide quantitative information on labour market needs. While these sources are very rich in information and can be nationally representative, they can also have important limitations. In less advanced economies they may not be conducted on a regular basis because of their high cost. More generally, indicators they provide, such as occupation or qualification, are only proxies for understanding actual skills requirements, and may not by themselves convey enough specific information to reliably guide action. The skills needs associated with particular occupations vary with context, and change over time.

New sources of data on skills have potential to provide real time and detailed information on skills needs in a cost-effective way. Technological advances, digitalisation and Internet platforms have made it possible to collect very rich and big datasets, or so-called “big data” for many purposes. Data on the content of job advertisements has been collected systematically from online job postings in a range of countries, creating huge datasets containing detailed information on the requirements advertised. Information typically recorded includes the specific skills needs and skills-related indicators in advertisements such as job titles, and requirements for qualifications, certifications and experience, as well as other information about the vacancy such as the employer, the economic sector, the occupational category and the geographic location. Collection of data happens in real time, and, in contrast with surveys that require time for processing before publication, data derived from online job vacancies can be used almost immediately. The richness of information featured in these datasets has drawn considerable attention, and has underpinned many publications, both from academia and from international organisations.

However, it is important to take into account limitations with using vacancies data derived from the web as a basis for labour market information. The sample of vacancies collected may not be representative of all web-advertised vacancies, and is unlikely to be representative of all job vacancies because of differences in recruitment practices by occupation and industry. Higher level skills are more likely to be advertised online, especially in less advanced economies. Online job advertisements chiefly cover the formal economy, so skills needed for informal employment are likely to be under-represented.

Skills requirements posted in online job listings provide a window into the landscape of detailed skills requirements, but they do not provide full listings of skills required in the manner of an occupational competency standard. There are no consistent standards as to what skills should be included in an advertisement. Many skills may be omitted because they are taken for granted by the recruiter, while those included may have been identified by the employer as salient because they differentiate the job from similar jobs. Other potential issues inherent to real-time data arise from a lack of structure, imperfect information, and measurement errors, especially related to duplicate observations and advertisements in which the number of jobs available is unspecified. Legal and regulatory matters may also be an issue, especially in particular in relation to privacy policies, which also pose challenges to extending the use of big data in labour market analysis to data sources other than vacancies data.

Analyses of vacancies data do not directly substitute for the main existing sources of labour market data. They provide flow-related measures inclusive of churn, where conventional labour market analysis is often based on measures of stock or on direct measures of flow that seek to filter out churn. It remains to be seen to what extent vacancies-based big data analysis will be mainstreamed into national skills anticipation systems, the extent to which it will be a complement or a substitute for other parts of these systems, and the extent to which it could provide a

fast track into skills anticipation for less developed countries without established systems. The future role of big data in skills anticipation beyond the vacancies data domain also remains unclear.

Objectives

The goal of the workshop is to allow all participating agencies to exchange experiences and ideas in relation to the use of big data for identifying labour market information. These discussions will consider:

- the feasibility of the use of big data in this context
- the potential of big data in skills analysis and the limitations associated with it
- how to advance the agenda in this area, share good practices, as well as find solutions to commonly found issues
- how these methods can best be deployed in developing economies

Agenda

DAY ONE

08:30-09:00 Registration of participants (R2 South Entrance)

09:00-09:30 Opening session

- *Srinivas Reddy*, Chief of the Skills and Employability Branch, Employment Policy Department
- Tour de table

09:30-11:00 Session 1: Stock-taking: knowledge-sharing on the usage of big data – conceptual and technical aspects

The aim of this session is to share knowledge and experiences on conceptual and technical aspects of applying real-time data to skills analysis. Rather than focusing on specific countries or regions, the idea is to have a broad analysis and discussion detached from any particular national context.

Moderator: Srinivas Reddy (ILO)

- *Konstantinos Pouliakas and Jasper Van Loo (Cedefop): “European Online Job Vacancies Analysis”*
- *Anastasia Fetsi (European Training Foundation): “Big Data for Labour Market Intelligence - Shaping, applying and sustaining knowledge”*
- *Fabio Mercorio (University of Milan – Bicocca): “Can we use big data for skills anticipation and matching? The case of Online Job Vacancies”*

11:00-11:30 Tea/Coffee break (R3 Delegates bar)

11:30-12:30 Session 2 : Usage of big data for skills needs assessment and matching: learning from advanced countries’ experience

This session will enable participants to share experiences on the actual use of big data for the analysis of skills requirements on the labour market within the context of more advanced economies. It aims to give an insight into academic research, and into practical applications from other institutions.

Moderator: Bolormaa Tumurchudur-Klok (ILO)

- *Claudia Plaimauer (3S) : “Using big data and semantic technologies for identifying labour market information in Austria”*
- *Renier Van Gelooven (S-BB): “The use of big data for skills anticipation and matching in the Netherlands”*

12:30-14:00 Lunch break

14:00-16:00 Session 2 (ctd): Usage of big data for skills needs assessment and matching: learning from advanced countries’ experience

- *Gábor Kismihók (Leibniz Information Centre for Science and Technology): “Lessons learned from studies on education-labour market matching”*
- *Stefan Winzenried (Janzz Technology): “From Big to Smart Data - The misconception of useful/predictive skills collections from Big Data”*
- *Tony Bonen (Labour Market Information Council): “Enhancing Skills Data in Canada – Connecting “big data” with traditional sources of labour market information”*
- *Andy Durman (EMSI UK): “Bringing traditional sense to the big data craze”*

16:00-16:30 Tea/Coffee break (R3 Delegates bar)

16:30-17:30 Session 3: Usage of big data for emerging and developing economies

Collection and analysis online job-vacancy data seems to provide good coverage of advanced economies. Emerging and less developed economies are facing more difficulties in applying this kind of analysis. Nonetheless, interesting studies and applications are already available for some developing and emerging economies, and this session is meant to give an insight into these.

Moderator: Ana Podjanin (ILO)

- *Carlos Ospino Hernández (IADB): “How far can your skills take you? Understanding skill demand changes due to occupational shifts and the transferability of workers across occupations” (via Webex)*
- *Sukriti (LinkedIn India): “LinkedIn’s Economic Graph: Data-driven insights for policy-making”*

DAY TWO**09:00-10:00 Session 3 (ctd): Usage of big data for emerging and developing economies**

- *Hirokichi Katayama (UNESCO): "Using real-time big data for TVET policies and strategies: Cases on Malawi, Myanmar and Mongolia"*
- *Gary Gan (JobKred): "Using AI to match skills supply to skills demand for industry 4.0" (via Webex)*

10:00-10:30 Tea/Coffee break (R3 Delegates bar)**10:30-12:30 Session 4: Connecting the dots: using and combining big data with a specific focus**

This final thematic session of the workshop on big data aims to provide insights into the application of big data analysis not necessarily in specific countries, but for specific purposes. These may be within the context of a certain economic sector and/or of particular type of jobs or skills needs. It also aims to show how the analysis of big data can be combined with other sources to widen the questions it can address.

Moderator: Olga Strietska-Illina (ILO)

- *Olga Strietska-Illina, Bolormaa Tumurchudur-Klok, Cornelius Gregg and Ana Podjanin (ILO): "Sharing experiences in combining big data with other methods"*
- *Inna Grinis (LSE/Goldman Sachs): "The STEM requirements of "Non-STEM" jobs: Evidence from UK online vacancy postings"*
- *Dan Restuccia (Burning Glass Technologies): "No Longer Optional: Employer Demand for Digital Skills"*

12:30 Closing remarks: Way forward and lessons learned

For more information:

Skills and Employability Branch (SKILLS)

Employment and Policy Department

Email: empskills@ilo.org

www.ilo.org/skills