BIG DATA AND THE COMPLEXITY OF LABOUR MARKET POLICIES:

New Approaches in Regional and Local Labour Market Monitoring for Reducing Skills Mismatches

Results of the 10th Annual Meeting of the

European Network on Regional Labour Market Monitoring (EN RLMM)

on 15 October 2015 in Milan, Italy

This document is based on the presentations and working group results at the Annual Meeting, which are available at **www.regionallabourmarketmonitoring.net. The recording of the event can be viewed HERE.**

European Network on Regional Labour Market Monitoring (EN RLMM)
Institute for Economics, Labour and Culture (IWAK)
Centre of Goethe University Frankfurt am Main, Germany

Network Coordinator: Dr Christa Larsen
c.larsen@em.uni-frankfurt.de
+49 (0)69 798 22152

Network Manager: Sigrid Rand
s.rand@em.uni-frankfurt.de
+49 (0)69 798 25474

The emergence and continuous development of digital society has an impact on our daily lives, also the labour market. Consequently, this leads to very fast changes. This development was taken as a starting point for exploring:

- In how far digitalisation is relevant for the topics which we are working on in regional and local labour market monitoring
- In how far new data sources are used for conducting our monitoring processes.

The latter relates to the application of Big Data in regional and local labour market monitoring.

1. The EU-Framework for Digitalisation

Before looking in detail on skills requirements related to the digitalisation and the application of Big Data for labour market monitoring, we will consider the political and legal framework conditions which are provided by the European Commission. One of the major objectives of DG Connect is to install a framework to support the development of the Digital Single Market in Europe (presentation by Lucilla Sioli): "A digital single market means an area where the free movement of goods, persons, services and capital is ensured and where citizens and businesses can therefore seamlessly access and exercise online activities under conditions of fair competition, irrespective of their nationality of place of residence". For facilitating the development of a digital market the EU Commission is active in three areas:

- Better access of consumers and businesses to digital goods and services across Europe (regulations on consumer and protecting co-operations)
- Creating the right conditions which enable digital networks and services to flourish (legislative approaches to transforming the current telecoms rules and the Audiovisual Media Services Directive as well as establishing a cybersecurity contractual public private- partnership)
- Maximising the growth potential of the Digital Economy (initiatives on data ownership, free flow of data and a European Cloud).

These framework conditions should support the fast development of a Digital Single Market.

The digitalisation will encompass all spheres of our economies. Work will be (and already has been in some areas) strongly reshaped. Therefore, it is of high importance that there is adequate information on the skills needed. According to the Digital Agenda Scoreboard of DG Connect measuring the impact of the digitalisation on the labour force shows the following: already 90% of all jobs require some level of digital skills and all jobs will change, whilst some may disappear. Today one third of the EU workforce has insufficient digital skills and 40% of companies trying to recruit ICT professionals have difficulty to do so. For 2020, 800,000 vacancies for IT professionals are estimated across the EU. This data show that for regional and local labour market monitoring and the regional and local labour market observatories it is of high importance to monitor the changes in jobs related to digital skills and the appearance of new digital jobs. This is crucial information especially for all actors and organisations involved in VET

Network Coordinator: Dr Christa Larsen
c.larsen@em.uni-frankfurt.de
+49 (0)69 798 22152

Network Manager: Sigrid Rand
s.rand@em.uni-frankfurt.de
+49 (0)69 798 25474

as well as in all activities for further education/training and life-long-learning. The Digital Agenda Scoreboard of DG Connect is a good example for accessing such skills requirements. This example gives general orientation for the regional and local labour market observatories but there is a need to explore further what the modes for the observatories having limited resources will be (money, time, IT-knowledge) in providing such information. One way is to use expert knowledge to learn fast from the demands of the enterprises in the region or the locality. However, there might be a new data resource which could even provide real time information: Big Data is data generated on the internet. How can it be used and exploited for regional and local labour market monitoring?

2. Big Data and Its Application in Regional and Local Labour Market Monitoring

There are two major understandings of Big Data: firstly, Big Data is data which is extracted directly from the Internet. Secondly, huge data stocks originating from different sources that have to be connected with each other (stored in data warehouses) can be denoted Big Data. In the case of the latter, the focus is on data generated in processes of public administration (routine data) and survey data and this focus will not be pursued further. Instead, we focus on the first definition, which involves Big Data from the Internet. This data can be exploited to not only describe the changed requirements in jobs and qualifications related to digitalisation, but to even provide insights in the behaviour of unemployed and employees.

One indication for the relevance of Big Data for Labour Market Monitoring is that there is a discussion in Cedefop on the role of Big Data for the framework of skills analysis (presentation by Vladimir Kvetan). Also, the CEDEFOP supports a project to develop a pan-European tool to exploit internet-based job-portals for providing real time (nowcast) information on skills demands of companies. The announcement that "focussing rather on skills intelligence (based on real-time data) than on forecasting elements" could give major orientation on the use and application of Big Data following the zenith of the forecasting times.

Big Data from the Internet could be a source for regional and local labour market monitoring. There are different sources to be distinguished. Firstly, there is a considerable amount of data in job portals. Secondly, there is the provision of meta-data which is generated by providers like Google. This is Google Trends, for example. Thirdly, there is an immense data stock which is generated in the social networks, especially the professional networks like LinkedIn (presentation by Christa Larsen). Currently, there is most experience with exploiting the job portals for regional and local labour market monitoring (presentation by Emilio Colombo). There is a major difference in the processing of this data compared to "conventional" data, since it has a huge volume which requires strong IT tools for extraction, storage and analysis and the data is quite heterogeneous. This means that there are structured data as well as a lot of unstructured data like text. But also one major advantage is that these are real-time data which could be continuously updated without time lags and additional costs. However, extracting data from the Internet means also that the data loses the context which might be of high importance for the correct understanding of it and there are even more challenges like methodological issues concerning data quality as there is no information on the population and sampling. Not to forget that Big Data is often individual data and there are chances to connect data from different sources on

Network Coordinator: Dr Christa Larsen
c.larsen@em.uni-frankfurt.de
+49 (0)69 798 22152

Network Manager: Sigrid Rand
s.rand@em.uni-frankfurt.de
+49 (0)69 798 25474

an individual level. This raises questions on data ownership, data protection and a lot of related ethical and legal issues which are not solved yet.

2.1 Job portals as a Source for Big Data

There is an interesting example "WollyBi – the Italian Labour Market Digital Monitor" on how to use job portals for generating real time information (presentation by Mario Mezzanzanica). Making Big Data from job portals useful for decision-makers and stakeholders requires that these data have to be turned into knowledge. Knowledge is then an end product of a data-driven discovery. To exploit job portal data involves selecting the job portals first. There are different types of job portals. Most of the time the job portals run by the public labour administration covers medium- and low-qualified jobs and the commercialised, private portals include high-qualified jobs with a focus rather on IT, consulting, technology and finances. To get a broad spectrum of professions to be covered often means that one has to exploit two or even more job portals. In regions and localities with rather specialised economies like food processing, textile or even IT, it might be also reasonable to focus on specialised job portals. After the selection of the portal(s) the data needs to be extracted. In most cases, information technology is used here like robots or scrapers. Some observatories do not have such competences. In this case they either outsource this task or they are looking for an IT-partner from universities or the corporate world (sponsor). Up to now there are no legal restrictions on data crawling. However, some portals are not willing to cooperate when they realise that a scraping approach is being used.

Raw data needs to be processed to detect duplications, for example. As soon as the data has been prepared in a process of data mining, the data is related to classifications like codes for professions etc. Subsequently, the classified information can be used for analysis. However, having huge amounts of data requires also systematic data storage in data warehouses. For a lot of regional and local observatories this is also a challenge. Solutions for this could be co-operations with universities or other public data providers. The analysis can show what exact skills sets are demanded for a certain profession. It could also provide an overview on what types of jobs are offered and how this changes over time. The abovementioned skills sets can also be used as a starting point for comparing professions based on the skills needed for professional performance. This way, overlaps between professions can easily be detected. Generally, this could be extremely valuable information for career changes, for further education and training as well as for the job placement of discriminated groups in the labour market. The digitalisation also supports the development of interactive interfaces in data provision to the users. This means that they can do the analysis which they need for their specific field of responsibility. The WollyBi project demonstrates that there could be different interfaces for different user groups like workforce agencies, business associations and unions, government employment agencies and training organisations (WollyBi website).

2.2 Meta-data as a Source for Big Data

Up to now, there is not a lot of application of meta-data in regional and local labour market monitoring. The single approach is applying Google Trends data (presentation by Andrea Fasulo). Google Trends

Network Coordinator: Dr Christa Larsen
c.larsen@em.uni-frankfurt.de
+49 (0)69 798 22152

Network Manager: Sigrid Rand
s.rand@em.uni-frankfurt.de
+49 (0)69 798 25474

data is a count of Google searches on a specific term like unemployment or job search. The count can be connected with counts on terms which are often connected with the first term. Additionally, the counts on specific terms could be selected for specific territorial units and they are provided as times series. There is a use of Google Trends by Network Members located in the National Statistical Office in Italy. They use these data and explore in how far this approach delivers reliable results predicting unemployment, for example. The current results show that the predictive power of Google Trends data is not worse than the one of conventional data. Therefore, there is the demand to further experiment with such data resources. It can be expected that meta-data could be a valuable, real-time and costless data source for regional and local labour market observatories as soon as their application is a little further explored. We ask the Network Members and the representatives of the regional and local labour market observatories to start collecting experiences with meta-data and feed this back to the European Network on Regional Labour Market Monitoring and its Members.

2.3 Social Media as a Source for Big Data

Up to now, the application of social media in regional and local labour market monitoring is only marginal. At this point, social media information is only used for the recruiting process. It is known that information form LinkedIn and Facebook is used in recruiting processes next to the ordinary application materials to get a deeper insight into the behaviour and self-understanding of the potential candidates. Often this information is also used in checking the validity of the information made in the formal application through looking at the network members of a person on LinkedIn. Up to now there is no information on how these issues affect the behaviour of potential applicants. All of this is very important background information for understanding the matching of demand and supply. It will take some effort and some exploration to find out in how far information from social media could be used for regional and local labour market monitoring. For example, it could be explored which experts are available in a certain region and how relevant this could be for regional labour market development and vice versa. Also here we would like to ask the Members of the European Network on Regional Labour Market Monitoring as well as the representatives of the regional and local labour market observatories to explore and experiment with social media information related to its application in regional and local labour market monitoring and feed this experience back into the Network.

2.4 Combining Different Sources of Big Data

In the commercial sphere, the example of Monster shows how different sources for Big Data can be combined to generate comprehensive pictures (presentation by Pietro Manenti). Developing and selling information on (potential) job applicants for companies/HR departments is a business field and the market seems to be growing. The framework for such an approach is improving the matching of supply and demand. Looking at this approach questions on data ownership, individual rights from a legal side as well as from an ethical point of view have to be raised. Although this example shows that there the technical tools for combining different data sources exist. For the Members of the European Network on Regional Labour Market Monitoring this example addresses the perspective in utilising Big Data from multiple sources. To do so, accompanying legal, ethical and foremost methodological issues need

Network Coordinator: Dr Christa Larsen
c.larsen@em.uni-frankfurt.de
+49 (0)69 798 22152

Network Manager: Sigrid Rand
s.rand@em.uni-frankfurt.de
+49 (0)69 798 25474

to be further explored. To support this, a Permanent Big Data Working Group was established within the European Network on Regional Labour Market Monitoring.

3. State of the Art, Challenges and Perspective of Big Data Application in Regional and Local Labour Market Monitoring

The usage of Big Data in Labour Market Monitoring is still in its infancy, but in the coming years, significant developments in the field can be expected which need to be followed closely.

For the most Network Members, working with Big Data is different from working with "conventional" data. Firstly, it requires a lot of specific knowledge and considerable IT competences. Also legal knowledge and methodological topics are very relevant. In practice, this means working in interdisciplinary teams. Currently, this is not the case in most labour market observatories. Therefore, strategic co-operations with universities and expert organisations need to be built. A second challenge is related to financial and technical resources. Data scraping, using robots and data storage in data warehouses as well as the creation of interactive web-based interfaces needs a lot of resources which are not available in regional and local labour market observatories. Also here, strategic co-operations which also could include finding sponsors in the corporate world or getting sponsored by public funding (programmes). Up to now, there is not a lot of support especially form public funds to be expected but the Network as well its members will have to work continuously on sensitising decision-makers in politics and public bodies for the relevance of this topic. If there is not a change in the funding in the near future, the developments will be so fast that the commercial actors will be the major information providers for all labour market actors. And the public labour market observatories, especially in regions and localities, could run out of business very fast.

The European Network on Regional Labour Market Monitoring will try to support the existence of this "neutral" public structure of nearly 600 regional and local labour market observatories across European regions (connected by the Initiative for Networking Regional and Local Labour market Observatories Across Europe), by founding a Permanent Working Group on Big Data and informing interested stakeholders on the state of the art. Also the Network will stay with the topic in 2016 by working on the effects of digitalisation on labour and working conditions. Some of the issues addressed in this documentation will be continuously covered.

4. The Permanent Working Group on Big Data

During the Annual Meeting, a Permanent Working Group on Big Data was founded. The group follows different objectives like observing the developments made by applying Big Data in labour market monitoring, supporting the exchange of information on this within the Network and beyond, improving the knowledge and skills needed for applying Big Data, analysing and communicating the need for public sponsoring as well as building strategic co-operations with Cedefop and the European Commission and other stakeholders on the European level. To cover these objectives, four topics should be in the focus of the group:

-
-

Network Coordinator: Dr Christa Larsen
c.larsen@em.uni-frankfurt.de
+49 (0)69 798 22152

Network Manager: Sigrid Rand
s.rand@em.uni-frankfurt.de
+49 (0)69 798 25474

- Technical and quality issues (experts will focus on the technical activity, data extraction, automated tools, methodologies, quality assessment, IT-infrastructure, data analysis and data accessibility)
- Decision support (experts will follow the development of communication formats and web interfaces in the further process of digitalisation)
- Social aspects of Big Data will be in the focus (experts will research the impact on real life of people and the related ethical issues)
- Legal aspects will be addressed (experts will focus on data ownership, legal frameworks and standards and criteria how to apply them). It is planned that the working group will report to the Network Members once a year during the European Day which is connected with the Annual Meeting

5. 11th Annual Meeting 2016 in Rovaniemi (Lapland) on the Topic "Digital Revolution and Its Effects on Labour: Opportunities and Challenges for Regional and Local Labour Market Monitoring"

As digitalisation is such an important development for the field of work, the activities of the European Network on Regional Labour Market Monitoring will be addressing this topic in 2016. In the focus of interest will be the exchange and discussion on how the digital revolution will influence labour. In particular, the activities in the network will explore how work is conducted, the working conditions and the possible differences between various branches, qualification levels, professions and target groups. If possible, we will reflect on how the regional framework conditions can influence these processes. Based on a Call, an anthology will be published in August for the 11th Annual Meeting of the EN RLMM.

SAVE THE DATE

18 August 2016: 11th Annual Meeting of the EN RLMM in Rovaniemi (Lapland)

19 August 2016: 6th European Day of the Initiative for Networking Regional and Local Labour Market3 Observatories Rovaniemi (Lapland)

For further information please contact the EN RLMM Manager **Sigrid Rand**, s.rand@em.uni-frankfurt.de, +49 69 798 25474.