

OJA Forum | 9 May 2022

Online Event

Agenda

Session 1 from 9:00 h to 12:00 h (all times stated in CEST)

09:00 **Welcome**

09:25 **The Online Job Ad Analysis Cycle – From collection to analysis**

Johannes Müller – &effect

To structure the discussion around challenges and methods in analyzing online job ads, Johannes Müller from &effect introduces the OJA Cycle heuristic. It follows the steps that one specific online job ad would take: collection, enrichment, aggregation and analysis. In his presentation, Johannes Müller introduces the OJA cycle project and preliminary findings.

09:55 **Social talk and break**

10:20 **Understanding national OJA markets – the role of landscaping**

Jiri Branka – CEDEFOP

Before engaging with actual data gathering and data production, it is highly advisable to understand the overall environment of OJA providers. CEDEFOP has developed a methodology for examining the context, content and coverage of the OJA's providers in the country – the so-called Landscaping exercise. This talk describes the experience and knowledge gathered about OJA providers across Europe from two landscaping exercises carried out in 2017 and 2021. Jiri Branka also refers to the latest knowledge gathered about Germany.

10:40 **Big data for the labour market: Use cases of OJA data**

Julia Nania – Emsi Burning Glass

This talk presents a collection of applied research examples using data about online job advertisements. It discusses methodological challenges and recommendations for the analysis.

11:00 **Discussion: Where are online job ads published and how to turn them into structured data?**

Johannes Müller (&effect), Jiri Branka (CEDEFOP), Julia Nania (Emsi Burning Glass)

In this first discussion, we deal with questions which arise at the beginning of the OJA cycle. Where are online job ads published, and what implications does this have on the collection and analysis of OJAs?

11:45 **Reflection and networking**

Session 2 from 13:30 h to 16:00 h

13:30 **Welcome back**

13:35 **Creating a dynamic, data-driven profession-skills ontology**

Kasper Kok – Textkernel

Over the last years, Textkernel has leveraged their database of hundreds of millions of vacancies to build a profession-skills ontology, which enables understanding the relationship between any two roles in terms of overlapping skills and skill gaps. In this session, Kasper Kok will give an overview of the process through which this ontology was built. He will also discuss the various challenges that arise: extraction and normalization of skills and professions, dealing with ambiguity and granularity, and making sure that the coverage is up-to-date and comprehensive.

13:55 **Using natural language processing to identify transversal skills and their labour market returns**

Fabio Manca – OECD

This presentation will show the OECD approach to identify transversal skills using an unsupervised Machine Learning approach. It will also give inputs into how this information can be used to assess the labour market returns associated with transversal skills.

14:15 **Break**

14:25 **Building an open jobs observatory for the UK**
Cath Sleeman – Nesta

Cath Sleeman will provide an overview of the journey to building an Open Jobs Observatory for the UK. The Observatory is a piece of data infrastructure in which Nesta automatically collects and processes online job adverts (with the permission of job boards). The processing includes the extraction of skills that were mentioned within the descriptions of adverts. In this presentation, Cath Sleeman will briefly describe the need for the Observatory, its progress to date, and the plans for the year ahead.

14:45 **Discussion: How to turn plain text into valuable information?**
Johannes Müller (&effect), Kasper Kok and Panos Alexopoulos (Textkernel), Fabio Manca (OECD), Cath Sleeman (Nesta)

In this second discussion, we deal with questions on data enrichment. In which ways can concepts be extracted from plain text and be translated into practical, actionable categories?

15:30 **Reflection: Online Job Ad Analysis Cycle**

15:45 **Feedback and farewell**