

OJA Forum

13. / 14. September 2022

Veranstaltungsort Bertelsmann Stiftung Gütersloh
Carl-Bertelsmann-Straße 256
33311 Gütersloh

1. Tag | 13. September 2022

Welcome

- 11:45 Uhr **Ankunft**
- 12:00 Uhr **Mittagessen**
- 13:00 Uhr **Begrüßung**

Session 1 Online-Stellenanzeigen als Forschungsdaten und deren Annotation - Erfahrungen und Best Practises

- 13:30 Uhr **Die Vorzüge einer öffentlichen Forschungsdaten-Infrastruktur für die Analyse von Stelleninseraten**
Dr. Felix Busch, Stellenmarkt Monitor Schweiz, Universität Zürich

Die empirische Sozialforschung ist auf eine hohe Verfügbarkeit von Daten angewiesen. Speziell für die Forschung etablierte Dateninfrastrukturen haben sich immer wieder als Katalysator für wissenschaftliche Innovation erwiesen. Dass ein einfacher und transparenter Zugang zu hoch-qualitativen Daten der Forschung mit Stelleninseraten noch fehlt, wird aus Publikationsverläufen und dem Austausch mit Forschenden sehr deutlich. In diesem Input spricht Dr. Felix Busch über die Vorzüge einer öffentlichen Forschungsdaten-Infrastruktur für Stelleninserate, etwa gegenüber der Nutzung von Daten privater Anbieter. Dabei zieht er Lehren aus der eigenen Datenakquise und der Zusammenarbeit mit gewerblich erhaltenen Datenquellen. Als Ausblick spricht Dr. Felix Busch darüber, was er ändern würde, wenn am Stellenmarktmonitor-Schweiz die Datensammlung neu aufgesetzt würde.

- 14:00 Uhr **Annotationsguidelines für deutschsprachige Stellenanzeigen: Eine Bestandsaufnahme inhaltlicher Herausforderungen**
Timo Schnepf, Bundesinstitut für Berufsbildung

Annotationsguidelines erlauben in Stellenanzeigen das konsistente Labeln relevanter Entitäten, wie bspw. jene mit Kompetenzbezug. Die Festlegung, wie eine Entität semantisch modelliert werden soll und anhand welcher Regeln sie

im Text zu markieren ist, erweist sich in der Praxis jedoch als nicht trivial. Entscheidungen, wie Entitäten in Stellenanzeigen annotiert werden, haben einen erheblichen Einfluss auf die Erstellung und Performance von Extraktionsmodellen, sowie die Aussagen von Analysen, die auf solchen Modellen basieren. Somit kommen Annotationsguidelines in Stellenanzeigenanalysen eine fundamentale Rolle zu. Der Vortrag greift relevante inhaltlichen Herausforderungen auf, erläutert sie an Beispielen und diskutiert mögliche Lösungswege.

14:30 Uhr **Annotation von Online-Stellenanzeigen – Erfahrungen und Best Practices**

Dr. Armin Mertens, IW Köln

Qualitativ hochwertige Trainingsdaten sind essenziell für das Gelingen von Machine Learning Projekten. Doch bei der Erstellung eigener, annotierter Trainingsdatensätze können viele Fehler gemacht werden. Deshalb werden hier Erfahrungen und Best Practices bei der Annotation von Textdaten (Online-Stellenanzeigen) vorgestellt. Dabei wird schwerpunktmäßig das Planen von Annotationsprojekten, die Auswahl geeigneter Tools und die Evaluation der annotierten Daten diskutiert

15:00 Uhr **Kaffeepause**

15:30 Uhr **Diskussion und Austausch zum Thema Online-Stellenanzeigen als Forschungsdaten und deren Annotation**

Session 2 Aktuelle Forschung Teil 1

16:30 Uhr **Befristungen in der Statistik der gemeldeten Arbeitsstellen: Analysen mit Text Mining**

Arsen Çelikel, Bundesagentur für Arbeit

Die Statistik der Bundesagentur für Arbeit (BA) erprobt in Kooperation mit dem IT-Systemhaus der BA erstmals Methoden des Text Mining für die Analyse von Stellenanzeigentexten mit dem Ziel, die amtliche Statistik der Gemeldeten Arbeitsstellen zu prüfen und ggf. anzureichern. Der Vortrag stellt die angewandte Methodik vor, mit der die Validität des Merkmals „Befristung“ unter Einbeziehung der Stellenanzeigentexten untersucht wurde, zeigt die erzielten Ergebnisse auf und gibt einen Ausblick, wie die gewonnenen Erkenntnisse in die statistische Verarbeitung miteinbezogen werden können, vgl. auch den im August 2021 veröffentlichten Methodenbericht.

17:00 Uhr **Kultur & Diversität in OJA - Profilanalyse & Handlungsstrategie für einen Mittelständler!**

Prof. Dr. Marc R.H. Roedenbeck - Technische Hochschule Wildau

Eine nicht zielgruppengerechte Sprache in Stellenanzeigen kann dazu führen, dass sich geeignete Jobsuchende nicht angesprochen fühlen, z.B. Frauen und Lebensältere. Basierend auf OJA eines Mittelständlers werden mittels des Linguistic Inquiry and Word Count (LIWC) die Gender- und Altersstereotypen sowie das Wertesystem (Unternehmenskultur) untersucht. Dazu werden ins

Deutsche übersetzte internationale Wörterbücher genutzt. Erste Ergebnisse zeigen eine selektive Eignung dieser. Die vorgeschlagenen Handlungsstrategien zur Optimierung des praktischen Nutzen sind a) die Überarbeitung der Wörterbücher für einen lokalen Sprachansatz, b) die Optimierung des Algorithmus sowie c) den Vergleich der Kultur-Ergebnisse mit der gelebten und avisierten Unternehmenskultur.

17:30 Uhr **Kurze Pause**

Session 3 Interaktive Session

17:45 Uhr **Interaktive Session unter anderem mit:**

- Dr. Fabian Stephany, Oxford Internet Institut/Alexander von Humboldt Institut für Internet und Gesellschaft
- Tonio Kellner, Textkernel

ab 19:30 Uhr **Abendessen im Restaurant La Trattoria**

2. Tag | 14. September 2022

09:00 Uhr **Begrüßung**

Session 4 Tools zur Visualisierung und Analyse von Online-Stellenanzeigen

09:15 Uhr **Visualisierung regionaler Analysen von OJA-Daten mit WordCloud_Mapper**

Gabriel da Silva Zech, Bertelsmann Stiftung / Hertie School of Governance

Das neu entwickelte Python-Package [WordCloud_Mapper](#) hilft Ihnen regionale Unterschiede in Ihren OJA-Analysen auf eine intuitive und zugängliche Weise zu kommunizieren. Durch die Erstellung von Wortwolken, die wie Regionen auf einer Karte geformt sind, ermöglicht WordCloud_Mapper z.B. die meistgefragten Berufe oder Kompetenzen nach Städten, Gemeinden, Bundesländern oder sogar Ländern zu visualisieren. In dieser Präsentation wird Gabriel da Silva Zech (Entwickler des Packages) in die grundlegenden Funktionen des Visualisierungstool einführen und in einer Live-Coding-Session zeigen, wie Sie schnell und unkompliziert Ihre eigenen Wortwolkenkarten erstellen können.

09:45 Uhr **Der Job Ad Decoder "JADE" - Dimensionen, Nutzungserfahrungen und Weiterentwicklung**

Univ.-Prof. Dr. Julia Brandl und Dr. Petra Eggenhofer-Rehart, Universität Innsbruck

Der Job Ad Decoder "JADE" basiert auf Präferenzbewertungen potenzieller Jobsuchender hinsichtlich Anforderungen, Benefits, Arbeitgeber- und Tätigkeitsbeschreibungen. Er beinhaltet die Dimensionen Gender und Alter (sowie demnächst auch Migrationshintergrund) und bietet empirisch geprüfte neutrale Alternativformulierungen an. Nutzer:innen-Erfahrungen bestätigen die Vorzüge des Tools und zeigen die Wichtigkeit eines verantwortungsvollen Umgangs kontextsensitiver Nutzer:innen mit Algorithmen.

10:15 Uhr **Ein Jobmonitor mit Online-Stellenanzeigen – erste Einblicke**
Gunvald Herdin, Bertelsmann Stiftung

Die Bertelsmann Stiftung wird in diesem Jahr einen „Jobmonitor“ veröffentlichen. Dabei handelt es sich um ein Dashboard mit regionalen Analysen aus Online-Stellenanzeigen. Wir möchten erste Einblicke in den Jobmonitor geben.

10:45 Uhr **Kaffeepause**

11:15 Uhr **Diskussion und Austausch zu Visualisierungen und Analysen von Online-Stellenanzeigen**

12:15 Uhr **Mittagessen**

Session 5 Aktuelle Forschung Teil 2

13:15 Uhr **Nachfrage nach transversalen Kompetenzen auf Basis von ESCO v1.1**

Dr. Martin Noack, Bertelsmann Stiftung

Die Anfang des Jahres von der EU-Kommission veröffentlichte ESCO v1.1 enthält eine eigene Taxonomie transversaler (überfachlicher) Kompetenzen mit 96 Kompetenzen in 24 Clustern von 6 Kategorien. Für die Extraktion aus deutschen Stellenanzeigen wurde auf ihrer Basis ein Wörterbuch mit 16.176 Suchworten entwickelt und auf 38 Mio. Stellenanzeigen von 2018-2021 angewandt. Der Vortrag stellt Unterschiede im Zeitverlauf nach Regionen (Kreisebene) und Branchen (KldB-Berufshauptgruppen) vor.

13:45 Uhr **Identifikation von Weiterbildungsmaßnahmen unter Berücksichtigung von OJA Daten**

Dr. Matthias Oschinski, University of Toronto / Belongnomics

Mit Förderung der Provinz Ontario (Kanada) wird derzeit ein digitales Modul entwickelt, das es Arbeitnehmer:innen erleichtern soll, zukunftsfeste Berufe zu identifizieren. Teil dieses Projektes ist der Entwurf eines Prototyps für eine Trainingsdatenbank. Diese soll eine gezieltere Auswahl geeigneter Kurse ermöglichen, um bestehende Fähigkeitslücken zu schließen. Bei der Entwicklung dieses Projektes wird mit generellen kanadischen Arbeitsmarktdaten, der O*NET Skills Taxonomy und Online Job Angebotsdaten des kanadischen Labour Market Information Council gearbeitet.

14:15 Uhr **Künstliche Intelligenz im europäischen Arbeitsmarkt**

Julia Nania, Lightcast

Künstliche Intelligenz, eine sogenannte Schlüsseltechnologie, ist einer der wichtigsten Faktoren im aktuellen Wandel des Arbeitsmarktes. Allerdings ist es schwierig das genaue Ausmaß der Nutzung zu bestimmen. Online-Stellenanzeigen bieten eine wertvolle Alternative zu Umfragen zur Verbreitung und Anwendung künstlicher Intelligenz im Arbeitsmarkt, welche Branchen besonders intensiv damit arbeiten, sowie welche spezifischen Kenntnisse benötigt werden. Unsere Analyse zeigt, dass einige europäische Länder im internationalen Vergleich leicht hinterherhinken. Dank der breiten Einsatzmöglichkeiten von KI verändert sich das Bild jedoch je nach Wirtschaftszweig und Land.

14:45 Uhr **Kaffeepause**

Session 6 Ethische und Juristische Perspektive

15:30 Uhr **Der Computer spielt Personaler:in? Chancen und Risiken algorithmischer Jobempfehlungen und des Robo Recruitings**
Teresa Staiger, Bertelsmann Stiftung

Der Einsatz algorithmischer Systeme und die damit verbundenen Potentiale sind mittlerweile in der Personalgewinnung weit verbreitet. Fallbeispiele wie die Benachteiligung weiblicher Bewerberinnen bei Amazon zeigen aber auch die Herausforderungen, vor die uns diese Systeme stellen. Denn auch in diesem Bereich setzen sich oftmals Diskriminierungen aus dem Analogen im Digitalen fort. Doch das muss nicht so sein: mithilfe der *Algo.Rules* können algorithmische Systeme so gestaltet werden, dass ihre Risiken verringert werden. Teresa Staiger, Projektmanagerin im Programm "Digitalisierung und Gemeinwohl", wird die *Algo.Rules* im Kontext des Robo Recruitings und der algorithmischen Jobempfehlungen vorstellen.

16:00 Uhr **Scraping als Wettbewerbsverstoß und Schutzfähigkeit von Algorithmen**
Dr. Sebastian Meyer, LL.M., BRANDI Rechtsanwälte Partnerschaft

16:30 Uhr **Feedback zur Veranstaltung**

16:45 Uhr **Ende des Forums**