scope

Contemporary Research Topics

work-based learning 4:
Technology

November 2022

https://doi.org/10.34074/scop.6004003

IN-DEMAND JOB SKILLS MONITOR

David Rozado, Francisco Rosas, Yoseob Shim

and Michael Holtz

Published by Otago Polytechnic Press. Otago Polytechnic Ltd is a subsidiary of Te Pūkenga, New Zealand Institute of Skills and Technology.

Article https://doi.org/10.34074/scop.6004003

IN-DEMAND JOB SKILLS MONITOR

David Rozado, Francisco Rosas, Yoseob Shim and Michael Holtz

ABSTRACT

A competitive economy needs a tertiary education system capable of equipping present and future members of the workforce with the skills needed by industry. Currently, the task of identifying what skills to teach within tertiary degree programmes is mostly carried out by regulatory agencies and educators who choose, to the best of their knowledge, the components to embed in degree courses based on their expertise and subjective intuition about what skills will better prepare learners for the job market. Could this model be suboptimal? A number of biases such as convenience sampling, motivated reasoning, group thinking, or incomplete information could creep in while academics and regulators engage in the curriculum design of tertiary education courses. We propose a complementary automated tool that quantifies the skills employers demand by monitoring and analysing the content of job advertisements within a given geographic region and aggregating them by job type. Such a system can empower tertiary education institutions to fine-tune the content of their courses accurately to meet actual, rather than guessed, industry demands for job skills.

INTRODUCTION

In a competitive job market, potential employers demand specific skills and prospective employees offer to supply that demand with their talent and abilities. In an efficient job market, the supply elastically adjusts to meet changing demand. This process is mutually beneficial for employers, employees and the wider economy. A mismatch between prospective employees' skills and industry demands can lead to job seekers' frustration (Altmann et al., 2018), lack of industry competitiveness (Middleton et al., 1993), decreases in productivity (Vandeplas & Thum-Thysen, 2019), and inefficient public spending in education and unemployment benefits (Cappelli, 1995). Thus, in order to thrive, a competitive economy needs a workforce equipped with the skills demanded by employers.

In modern societies, the task of equipping present and future members of the workforce with the skills needed for their career falls largely under the responsibility of tertiary education institutions: universities, polytechnics, institutes of technologies and other tertiary education providers (Salmi, 2003). Regulatory agencies and staff employed by tertiary education institutions often use their subject expertise and intuition to decide the content and skills to be taught/develop within degree programs. We contend that a number of cognitive biases such as motivated reasoning, convenience sampling, group-thinking or incomplete information can creep-in during this process (Vandeplas & Thum-Thysen, 2019). Furthermore, academics and regulators are often disconnected professionally from industry, with the majority of lecturers and professors spending most of their non-teaching time in research and departmental administration tasks that are often disconnected from industrial realities. Thus, we suggest a complementary data-driven tool to help tertiary education institutions determine the skills that are in-demand by industry.

The present work was motivated by previous findings suggesting a growing gap in the New Zealand job market between demands from employers for certain skills and an under-supply of qualified potential employees (NZTech, 2021). This mismatch could suggest under-investment in developing the existing New Zealand workforce and/or lack of coordination between industry and tertiary education providers. The dramatic skills shortages in some strategic industries (NZTech, 2021) drive a heavy reliance on immigration to supply unmet demand and create a national vulnerability in the receptor country and a brain drain in the source country.

Our proposal consists of an automated tool that monitors job advertisements within a geographic region (that is, city, state, country) and quantifies the prevalence of technical and professional skills for different professions in those job advertisements. The data is then aggregated by job type and conveniently visualised to provide an overview of the most demanded skills for a given profession. We contend that our tool could help tertiary education providers to fine-tune the content of their programmes to better meet current industry demands. Such a data-driven, implicit coordination between industry demands and the content of academic degrees could help mitigate professional skills gaps in the job market (APAC, 2021).

It is important to emphasise that our proposed system should not be considered a substitute for human judgment. We instead consider it a complementary tool that augments human expertise and decision-making by leveraging a quantitative overview of the skills demanded in a particular job market. We can think of several pedagogical reasons for which an educator might justifiably choose to prioritise the teaching of a tool or skill over another even when the former is less demanded by industry than the later. Thus, the ultimate curriculum design choices should still be made by regulatory bodies and educators who can integrate our quantitative data analytics with their contextually-aware expertise.

IN-DEMAND JOB SKILLS MONITOR

We briefly describe next the web application prototype we have created for illustration purposes. Our application quantifies the skills demanded by New Zealand industry by automatically monitoring job announcements in popular NZ employment advertisement sites and aggregating mentions of required skills by job type.

Conceptual overview

The *In-Demand Job Skills Monitor* web application consists of an online interface that allows a user to choose from a group of professions, a role or job title of interest in order to retrieve visualisations that quantify the technical and professional skills that are currently in-demand for such a role in the New Zealand job market.

The job types are classified within the application into several major categories such as *Information Technology* or *Medical & Health*. The major categories are, in turn, divided into common subcategories such as *Developers* or *System Administrators* and *Nursing* or *General Practitioners* respectively.

The application is comprised of two parts, a 'Backend' to store, process and serve data and a 'Frontend' which retrieves this data and transforms it into a visual representation.

Backend

The setup of the system consists of a Linux virtual machine on Google Cloud which holds a containerised Postgres database using Docker, an automated Python web scraping script and a FastApi GraphQL API server to expose the data to the outside world. An overview of the Backend component of the application is shown in Figure 1.

The job listings are scraped from popular job advertisement services, parsed and saved into the database. In case of dynamic content, the automation tool, Playwright, is used.

From the scraped information, a sample of job descriptions for a given job category is then used to manually extract job type-specific vocabulary comprising a set of terms for what the employers demand for this job type. This process allows us to define new classifications which will be used for data labelling using the annotation tool UBIAI and regular expressions. As a result, around 120 job descriptions are manually labelled and then used to train and fine-tune a ROBERTA transformer model for the task of Name Entity Recognition.

Once we have a trained language model, we use it to label the rest of the job descriptions for the job category in question. Then we count the number of mentions for each term in a category in all the job descriptions.

Lastly, we create a GraphQL API to query all the processed data in the database and serve it using Ariadne and FastApi to make it available to the Frontend.

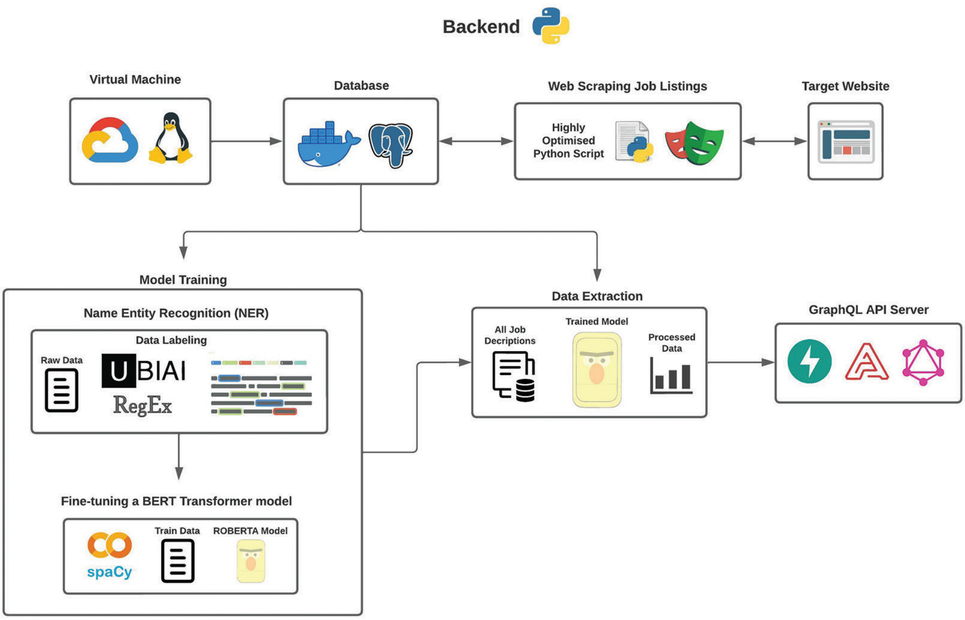

Figure 1. Conceptual overview of the Backend side of the In-Demand Job Skills Monitor App.

Frontend

A Static Site is generated with a NextJS application based on the Javascript framework React, which provides server-side functionality. This application is styled using the Material UI component library for React which abides by Google's design principles. An overview of the Frontend component of the application is shown in Figure 2.

This application retrieves the data using Apollo Client, a third-party library that provides an easy-to-use set of tools to fetch information from GraphQL APIs. Once all the necessary data is fetched, we use the libraries Mapbox and ApexCharts to help us visualise the information in the form of maps and charts, respectively. We apply animations using the React library Framer Motion and then proceed to deploy this application to Vercel, a platform to host web applications. Every time new data is scraped on the Backend, we redeploy the site to provide up-to-date information.

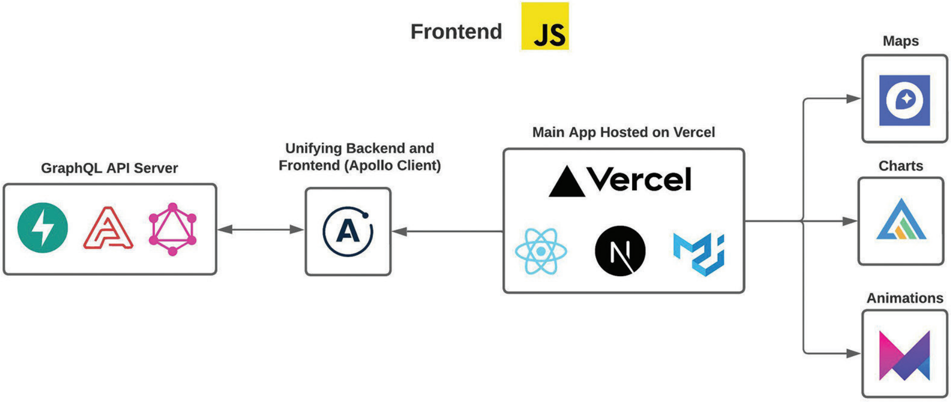

Figure 2. Conceptual overview of the Frontend side of the In-Demand Job Skills Monitor App.

Hardcoded skills vs machine estimated skills

Analysing natural language text requires efficient mechanisms to deal with ambiguity. When seeking information in text corpora, it is often helpful to consider 'context' and be cognisant that simple string matching through regular expressions is often insufficient for information retrieval purposes. That is, there are often an enormous amount of possible text configurations to express ideas such as the set of skills required for a job type. Machine learning-based natural language models fine-tuned for fuzzy information retrieval can help to identify such instances. This is often achieved by language models that represent the semantics/syntactic features and relationships between terms as a structured numerical vector that can be easily queried for fuzzy matching.

To illustrate this idea, let us say we want to match jobs that require *communication skills*. A simple template matching query system searching for the bigram *communication skills* would miss job listings that talk about *verbal skills* or *being good with words* unless such expressions are included explicitly as a set of synonyms for *communication skills*. Such an approach has obvious limitations for scaling. In vector space however, all semantically similar terms to *communication skills* would be located in adjacent regions in vector space by virtue of their semantic similarity. Thereby, a machine learning-powered language model is able to detect job types that require communication skills even if they all use different terms to express that concept.

Prototype

Figure 3 shows an illustrative output of the system for the query *software developer*. The figure clearly shows how our prototype application has captured the frameworks, programming languages, collaborative tools, database management systems and other technical tools that are most in demand by the New Zealand IT industry at the time of the query (November 2021) for the *software developer* profession. Our system has also captured the soft or professional skills that are often requested by potential employers for this role as well as the companies offering the most jobs in this category, the average salary and a geographic distribution of this job type around the country.

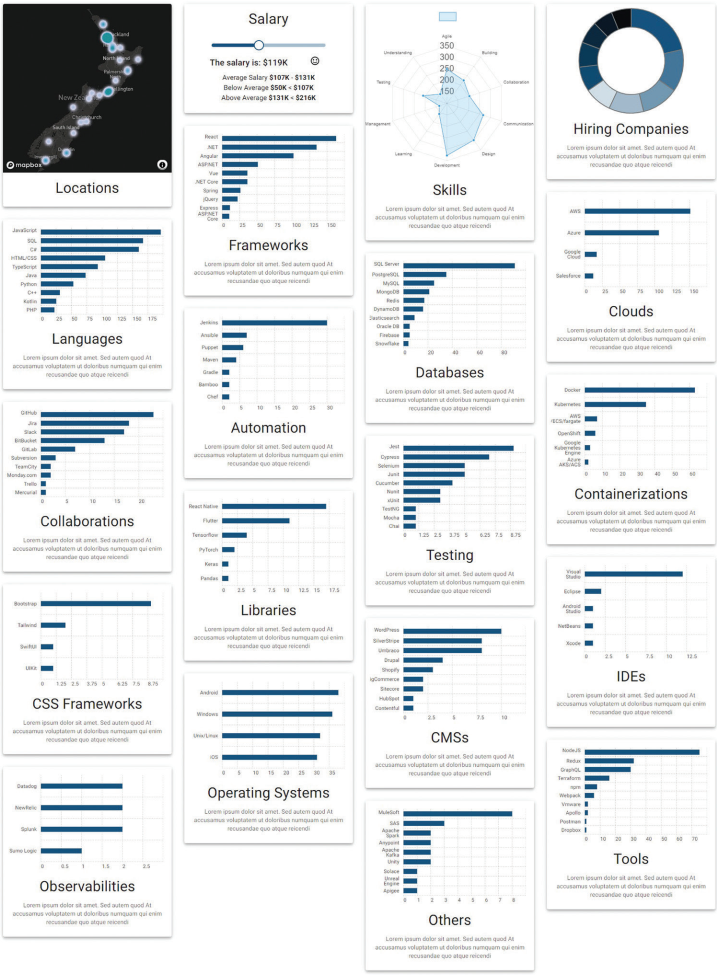

Figure 3. Visualisation of skills demanded by New Zealand employers for the role of 'software developer' as estimated from job ads in popular job advertisement sites.

CONCLUSION

This work has argued that the process of choosing content to be embedded in the educational programmes offered by tertiary education providers is often carried out in an ad-hoc manner that often leverages the expertise of regulators and educators to estimate the skills needed within a geographical job market. While this approach can be informative and valuable, it can also be limited by intrinsic cognitive biases and incompleteness of information available to academic staff and regulatory bodies. We propose a complementary system that estimates technical and professional skills demanded by industry within a given geographic region by monitoring job advertisements and quantifying the degree to which different technical and professional skills are demanded by job type. Our proposed system can help fine-tune the academic content delivered by educational institutions to actual, rather than guessed, skills demanded by industry. We hope our proposal and existing prototype can spark future investigations into how tertiary education providers can better fulfil industry demands for a skilled labour force to ensure the future employability of their students.

David Rozado is an Associate Professor at Otago Polytechnic in New Zealand. David holds a Ph.D. in Computer Science from the Autonomous University of Madrid. His research interests are computational content analysis and accessibility software.

https://orcid.org/0000-0001-6849-4746

Francisco Rosas is a graduate from Otago Polytechnic in New Zealand who holds a Bachelor of Information Technology. He has a passion for creating web and mobile applications that implement artificial intelligence.

Yoseob Shim is a graduate from Otago Polytechnic in New Zealand who holds a Bachelor of Information Technology.

Michael Holtz received his bachelor's degree in Information Science from the University of Otago in 1995. He worked as Technical Lead and Technical Consultant at Spark New Zealand. Currently, he is Head of IT programmes at Otago Polytechnic in New Zealand.

REFERENCES

Altmann, S., Falk, A., Jäger, S., & Zimmermann, F. (2018). Learning about job search: A field experiment with job seekers in Germany. *Journal of Public Economics, 164*(August), 33–49. https://doi.org/10.1016/j.jpubeco.2018.05.003

APAC Technical and Vocational Education Forum. (2021). "Speakers." Retrieved February 25, 2022, from https:///www.apactvetforum.com/speakers

Cappelli, P. (1995). Is the 'skills gap' really about attitudes? *California Management Review, 37*(4), 108–124. https://doi.org/10.2307/41165813

Middleton, J., Ziderman, A. & Adams, A. (1993). *Skills for productivity: Vocational education and training in developing countries*. Oxford University Press.

NZTech. (2021). *Digital skills Aotearoa: Digital skills for our digital future*. New Zealand Digital Skills Forum. Retrieved February 25, 2022, from https://nztech.org.nz/reports/digital-skills-for-our-digital-future/

Salmi, J. (2003). Constructing knowledge societies: New challenges for tertiary education. *Higher Education in Europe, 28*(1), 65–69. https://doi.org/10.1080/0379772032000110125

Vandeplas, A., & Thum-Thysen, A. (2019, July). *Skills mismatch and productivity in the EU*. Directorate General Economic and Financial Affairs, European Commission, 100. Retrieved February 25, 2022, from https://ideas.repec.org/p/euf/dispap/100.html