The Job Ads Database (JADB) Prototype: Building a New Resource for Historical Labour Market Research

Johansson, Mathias; Gladoić Håkansson, Peter; Karlsson, Tobias; La Mela, Matti

2026

Document Version:
Publisher's PDF, also known as Version of record

Lund Papers in Economic History

No. 264, 2026

The Job Ads Database (JADB) Prototype

Building a New Resource for Historical Labour Market Research

Mathias Johansson, Peter Gladoić Håkansson, Tobias Karlsson & Matti La Mela

DEPARTMENT OF ECONOMIC HISTORY, LUND UNIVERSITY

The Job Ads Database (JADB) Prototype

Building a New Resource for Historical Labour Market Research*

Mathias Johansson♣, Peter Gladoić Håkansson♦,

Tobias Karlsson♥ & Matti La Mela♠

Abstract

This paper introduces a novel dataset that serves as a prototype of what we call the 'Job Ads Database' (JADB). The underlying job advertisements are gathered from the Swedish digitalized newspaper corpus hosted by and made accessible through the Swedish National Library (Kungliga Biblioteket). JADB enables studying how the frequency and content of job advertisements change over time and sheds new light on the role of newspapers in recruitment, both relative to earlier practices and to public employment offices. We compare JADB data with official employment statistics of notified vacancies and applications from the Swedish Public Employment Administration (SPEA).

Keywords: job advertisements; newspapers; historical labour markets; recruitment practices; Sweden; digitized newspapers; HISCO; labour demand; nineteenth century; twentieth century

JEL-codes: N33, N34, J63, J64, C55, C81

---

* The work with setting up the JADB prototype has partly been funded by the Department of Society, Culture and Identity, Malmö University.

♣ Department of Arts and Cultural Sciences, Lund University, 0000-0002-3338-0551

♦ Department of Society, Culture and Identity, Malmö University, 0000-0003-0018-8720

♥ Department of Economic History, Lund University, 0000-0002-4778-6317

♠ Department of ALM, Uppsala University, 0000-0003-0340-9269

1 Introduction to JADB

Firms and other organizations have used newspapers to attract and hire workers since at least the nineteenth century. Yet we are only beginning to understand how this recruitment practice evolved. To address this gap, we are developing the Job Ads Database (JADB) prototype,[1] a new dataset based on job advertisements in Swedish newspapers. In this version we have limited data collection to one newspaper – the Stockholm-based *Dagens Nyheter* – for the period 1870–1970, but our intention is to scale up the dataset to cover all digitized newspapers from 1850 to 2000. JADB provides a new window into the historical labour market, allowing researchers to track changes in the demand for labour, the skills required for different occupations, and the evolution of recruitment practices over a century of significant economic and social change. More generally, with the digitisation of historical newspapers in different national contexts, job ads are receiving increased attention in economic history research, and computational approaches have been leveraged to extract information such as occupation titles for the analysis of historical labour markets (Schulz et al., 2014; Ros et al., 2020; Calabrese & Van Leeuwen, 2023; Gladoić Håkansson et al., 2023; Kleinert & Moelzer, 2024; Venglarova et al., 2024; Bi et al., 2025). To our knowledge, however, JADB is the first database that also allows the users also to explore the data through its public interface.

This paper focuses on the prototype of JADB and provides a detailed description of its construction. We discuss the data sources, the methods used to extract and classify job advertisements, and the structure of the resulting dataset. We also present some preliminary findings, focusing on long-term trends in job advertising and a comparison with data from the Swedish Public Employment Administration (SPEA). This paper is intended as a guide for researchers interested in using the JADB – as well as historical job advertisements more generally – in their own work.

[1] See https://JobAdsData.github.io/HISCOTrend/.

1.1 Theoretical framework – job matching and search

Although historical and contemporary examples of physical hiring fairs exist, the exchange of information between buyers and sellers of labour has, since the nineteenth century, largely been channelled by newspapers and other media. Job ads allowed employers and workers to learn about available applicants and vacant positions, respectively. In the ads, employers and workers can also begin the process of screening potential applicants and their skills (Deming & Kahn, 2018). The employer can, for example, require applicants to submit diplomas and references and the workers can signal what kind of employer they are looking for. Newspaper ads also reflect the geographical scope of labour markets and structural divisions between workers and jobs. These divisions may, for example, be based on gender, age, or occupation and can be seen in the light of theories of dual or segmented labour markets (cf. Leontaridi, 1998; Piore, 1979).

From the job seeker's perspective, search theory has examined which recruitment channels are most effective. Early studies by Holzer (1988) Holzer (1988) showed that informal methods like "friends and family" were highly effective both in terms of job offers received and search time. A substantial body of research has since investigated the importance of informal search channels and social networks (See among others: Battu et al., 2011; Bentolila et al., 2010; Bramoullé & Saint-Paul, 2010; Calvó-Armengol & Jackson, 2004; Calvó-Armengol & Zenou, 2005). This research demonstrates that informal channels can be rational and effective from the job seeker's perspective. However, the emergence of formal recruitment channels such as newspaper ads offered pathways that could reach beyond personal networks, potentially lowering search costs and expanding opportunities for both employers and job seekers.

Understanding how different recruitment channels emerge and evolve requires attention to technological, economic, and institutional change. The newspaper job ad, by expanding the reach of recruitment and facilitating matching between employers and workers, became increasingly relevant in the late nineteenth and early twentieth centuries. To better understand how these dynamics played out in practice, we now turn to a historical overview of recruitment practices and employment services.

1.2 A brief history of recruitment practices

Our knowledge of how recruitment practices have evolved over time is patchy. A useful point of departure is, however, the transition from a market where most employment contracts were annual to a market where contracts could be entered into and dissolved on shorter notice. This transition took place in societies that industrialized in the nineteenth century. Another important shift in the character of the labour market, including recruitment practices, was the transition into large-scale and bureaucratic organizations, often seen in the early twentieth century. This coincided with an increased involvement of the state in the operation of the labour market, with the introduction of public employment services.

Mobility in urban labour markets in the late nineteenth century was extensive (Fishback, 1998). Traditional recruitment channels, such as hiring fairs, became less reliable. It was in this context that newspapers' sections for job ads and recruitment agencies appeared and could play an increasingly significant role, although word-of-mouth and social networks remained influential (Licht, 1992; Kleinert & Moelzer, 2024; Bi et al., 2025; Rosenbloom, 1994).

In newspapers, both employers and job seekers could read and post their own ads. Table 1 presents the prices for a one-month subscription and for announcements in *Dagens Nyheter*'s Stockholm edition between 1890 and 1925. The announcement price refers to a three-line ad in the lowest price category[2] and is also shown in fixed (real) prices. We see that the announcement price increased both relative to the price of the newspaper and in real terms: between 1890 and 1925, the real price of our example announcement more than doubled. However, as we will see in a later section in this paper (Figure 2), the share of job ads rose over time, which would suggest that the ads became an increasingly important source of income for the newspapers, but also that the newspaper was a useful and important channel for recruitment.

[2] The price varies depending on the location in the newspaper, and announcements published on Saturdays cost an additional 10 %.

Table 1. Prices for Dagens Nyheter: subscription and announcement costs in nominal and 1929 fixed prices (Swedish kronor)

Year	DN one month subscription price	3-row announc. (lowest price category)	Announc. price/1-month subscription	Announc. price in 1929 fixed prices
1890	1.10	0.36	0.33	0.74
1895	1.10	0.42	0.38	0.92
1900	1.10	0.53	0.48	1.02
1905	1.10	0.60	0.55	1.15
1910	1.30	0.66	0.51	1.17
1914	1.30	0.66	0.51	1.10
1920	5.00	2.52	0.50	1.58
1925	3.50	1.98	0.57	1.88

Note: The prices refer to the Stockholm edition and for February (except for 1920, which refers to July, while no information available for earlier). Consumer Price Index: Ljungberg (2025).

Source: Svenska Tidningar, *https://tidningar.kb.se/*.

During the interwar period, state involvement in labour markets increased. Among other things, the emerging welfare states took responsibility for reducing unemployment by facilitating the intermediation of labour. Private employment agencies had existed since the nineteenth century but were criticised for profiting from the vulnerable position of unemployed workers.

In Sweden, the first public employment services started in 1902 in Helsingborg. Thereafter, local employment services spread over the country until it turned into a national Swedish Public Employment Administration (SPEA)[3] (Håkansson & Tovatt, 2017).

In 1910, a national list of vacant jobs (*riksvakanslistan*) appeared. Initially, it mostly reported vacancies concerning domestic service and agricultural work. Starting in the 1930s, the national vacancy list was broadcast on the radio every Thursday as a spoken job announcement. This practice continued until the 1970s. In the 1960s and 1970s, public employment agencies in larger cities also offered a service called *Fröken Plats*, where job seekers could use a telephone to dial a number and hear job vacancies read aloud. This service was particularly important to mediate temporary jobs to housewives (Arbetsförmedlingen, n.d.).

SPEA began publishing job ads in its own journal, *Platsjournalen,* in 1963. Initially, these ads did not disclose information about the employer. Job seekers were required to contact the

[3] first in the form of *Statens Arbetsmarknadskommission* and from 1948 as *Arbetsmarknadsverket*

employment services for further details. This changed in the 1970s, when job ads in *Platsjournalen* began to include the employer's name.

2 Data and Dataset construction

The JADB is constructed from the collection of digitized historical newspapers hosted by the Swedish National Library (Kungliga Biblioteket). The newspapers are not segmented into text units (like article or advertisement), and we therefore must search for the job ads indirectly, through keywords such as job titles together with markers for job ads. As a result, we can only measure the number of page hits, not individual ads or advertised positions.

2.1 HISCO

The Historical International Standard Classification of Occupations (HISCO) is a comprehensive, hierarchical classification and list of occupational titles, containing spelling variations, standard spellings, and genders (van Leeuwen et al., 2002). A Swedish version has been assembled by the research infrastructure SwedPop (Edvinsson & Westberg, 2024; SwedPop, 2022) containing 179,326 different spellings and forms of occupation titles. After removing non-occupational titles like "widow" and "unemployed", a list of 150,554 'actual' occupations remains. This list that we use when querying the database.

Table 2 shows the distribution of spelling variations and 'standardized' titles form the HISCO dataset, with the non-occupational titles removed. This distribution is clearly skewed toward male occupational titles, which make up more than two thirds of all the spelling variations and standardized occupational titles. Women's titles are in second place, and titles classified as being applied to both genders at a distant third place. However, what is not evident from this table is how the various spelling variations from different gender groups relate to the same 'standardized title'.

In Table 3, we have recalibrated the gender division, moving all spellings that share a standardized title with the other gender, the 'both'-category are all moved to this joint category. For example, *barberiarketare* (masculine) and *barberiarbeterska* (feminine) are both mapped to the standard title *barberarearbetare* – barber – even though they are gendered titles. In this redistribution, the spellings of men-only titles have been reduced to less than half the total

whereas women-only titles were reduced by ~15% and the non-exclusive titles are now up to nearly 1/3rd of all spellings. The distribution of standardized occupational titles, however, did not change significantly.

Table 2. Gender distribution of occupational titles and spellings in HISCO

Gender	Spellings	Occupations
Man	102,208	46,582
Woman	39,964	20,264
Both	8,382	6,025
Total	150,554	72,871

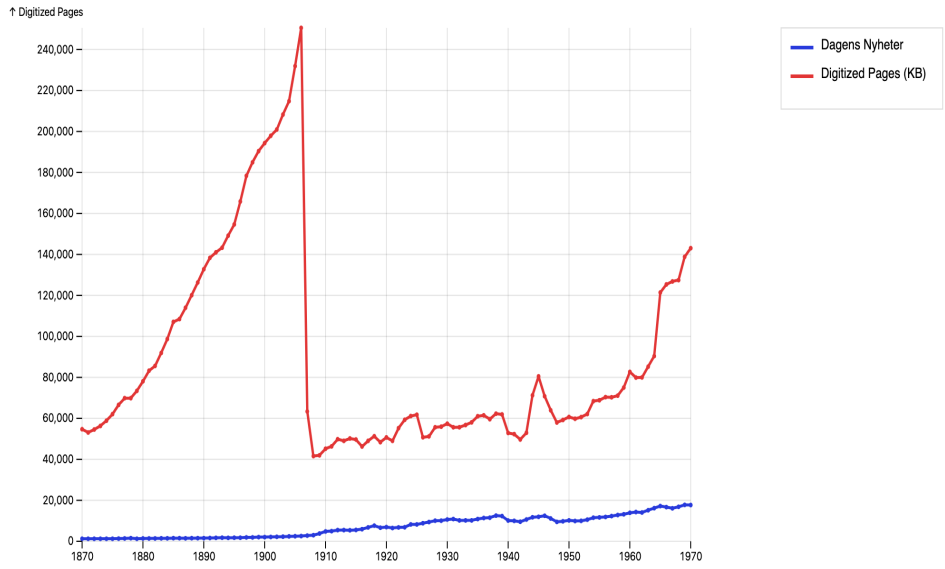
Table 3. Recalibrated gender distribution of occupational titles and spellings in HISCO

Gender	Spellings	Occupations
Man	71,123	43,343
Woman	33,469	18,851
Both	45,962	6,346
Total	150,554	68,530

2.2 The Swedish National Library's Newspaper Database

While the KB newspaper database is a rich source for historical research with nearly 2.5 million digitalized newspaper pages, it has many limitations. One limitation is the inconsistent coverage between 1906 and 2014 (Haffenden & Sikora, 2025; Karlsson, 2019). Figure 1 shows the distribution of the total number of pages in the entire newspaper corpus (red) and the total number of pages the corpus has of *Dagens Nyheter* (blue). KB reports full coverage up to and including 1906, as seen in Figure 1. After this peak the level of coverage is unclear, and the progress of current digitalization efforts are not made public. Nevertheless, it is clear that the overall coverage from 1907 is severely lacking for several decades afterwards. To ensure a consistent and comparable dataset across our entire study period we have chosen to focus on *Dagens Nyheter*, one of Sweden's major national newspapers. It was first published in 1864, was issued throughout the period 1870-1970, and contained job ads throughout.

Figure 1. Coverage of the Swedish National Library's newspaper database (blue) and digitalization of Dagens Nyheter (red) between 1870-1970



Note. The coverage is given as the number of pages.

Source: *Svenska Tidningar*, https://tidningar.kb.se/.

2.2.1 Accessing the newspaper corpus

There are two public access points into this corpus, first is the Graphical user interface offered up at "https://tidningar.kb.se", and an undocumented, unversioned API available at "https://data.kb.se/search/". For this study, we relied on the latter, as it gives easier programmatic access to run thousands of queries. Part of the response from the API is a timeline of the number of hits – the number of pages found each year for the query.

2.3 Indexing Job Advertisements

We extracted job advertisements using a combination of keyword searches and the HISCO list. The process was as follows:

1. **Occupation Titles**: We used all spelling variations of 'actual' occupation titles listed in the Historical International Standard Classification of Occupations (HISCO). For some examples related to messenger jobs, see Table 4.

Table 4. Examples of title spellings in HISCO

Occupational Title	Sex	Occupational Title, Standardised	HISCO
Springgosse i bokhandel	1	Bokhandelspringosse	37040
Springpojke på bokhand.	1	Bokhandelsspringpojke	37040
Bokhandelsspringpojke	1	Bokhandelsspringpojke	37040
Kaféspringflicka	2	Caféspringflicka	37040
Caféspringflicka	2	Caféspringflicka	37040

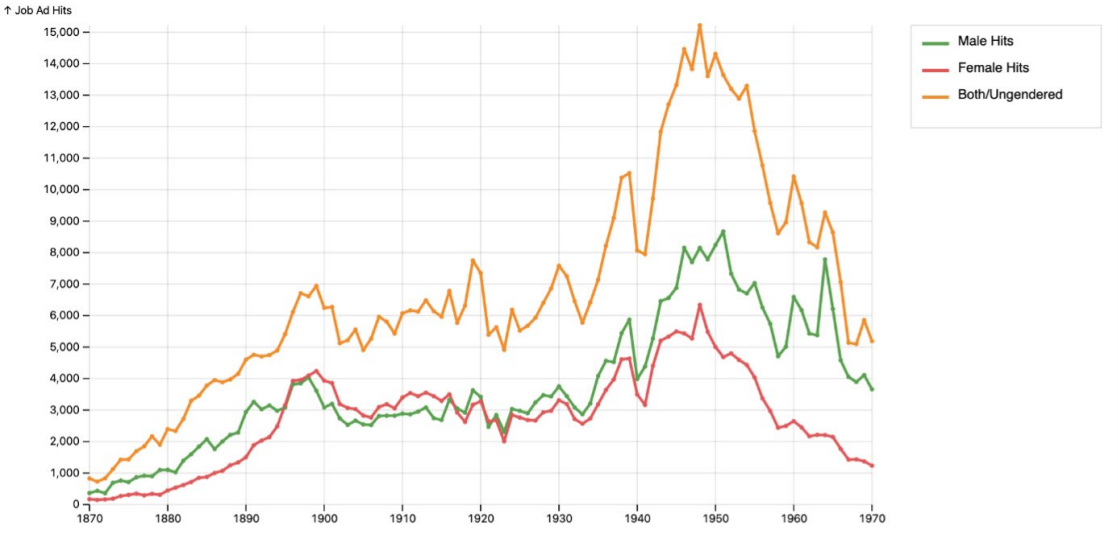
2. **Keyword Search**: We searched for each occupation title in the newspaper database using a specific search query. The query was designed to find pages where the occupation title appeared near a set of "marker" words that are commonly found in job ads. We used the markers "SÖKES" (WANTED), "PLATS" (POSITION/VACANCY), and "ERHÅLLES" (AVAILABLE). The search query was: `"<OCCUPATION> <marker>"~5`, which returns pages where the occupation title and marker are within 5 words of each other. For a more detailed an technical example (see Appendix A).
3. **Data Collection**: The searches were performed using the National Library's API (https://data.kb.se/search/docs)[4]. This API has a non-specified rate limit, which we tried to respect during the data collection process. We performed two sets of searches for each title: one using *all available newspapers*, and another one using only *Dagens Nyheter*. For this paper, we focus on the *Dagens Nyheter* data, as it is one of the few fully digitalized newspapers from the period. The search results are returned as the number of pages on which a match is found. As noted above, this means that we count the number of pages with at least one job ad for a specific occupation, not the total number of ads or advertised positions. While this method may underestimate the number of ads for very common occupations, we believe it provides a reliable measure of the long-term trends in job advertising (see Gladoić Håkansson et al., 2023).

[4] The routing of /docs is somewhat misleading as the name implies that it leads to documentation. Rather, it leads to an interface listing a lot of the endpoints and possible arguments. It also offers a great GUI for experimenting with the API. But there is nothing in the way of actual documentation.

2.4 Preliminary Findings

Figure 2 shows the number of page hits for male, female, and ungendered occupations in *Dagens Nyheter* from 1870 to 1970. Gender-neutral occupations account for a significant portion of the hits throughout the period. The number of hits for male occupations is generally higher than for female occupations, but both follow a similar trend, with peaks around 1899 and 1949. Towards the end of the period, the number of hits for female and ungendered occupations declines more rapidly than for male occupations, which may reflect broader shifts in the labour market.

Figure 2. Gendered-job title hits in Dagens Nyheter, absolute numbers 1870-1979



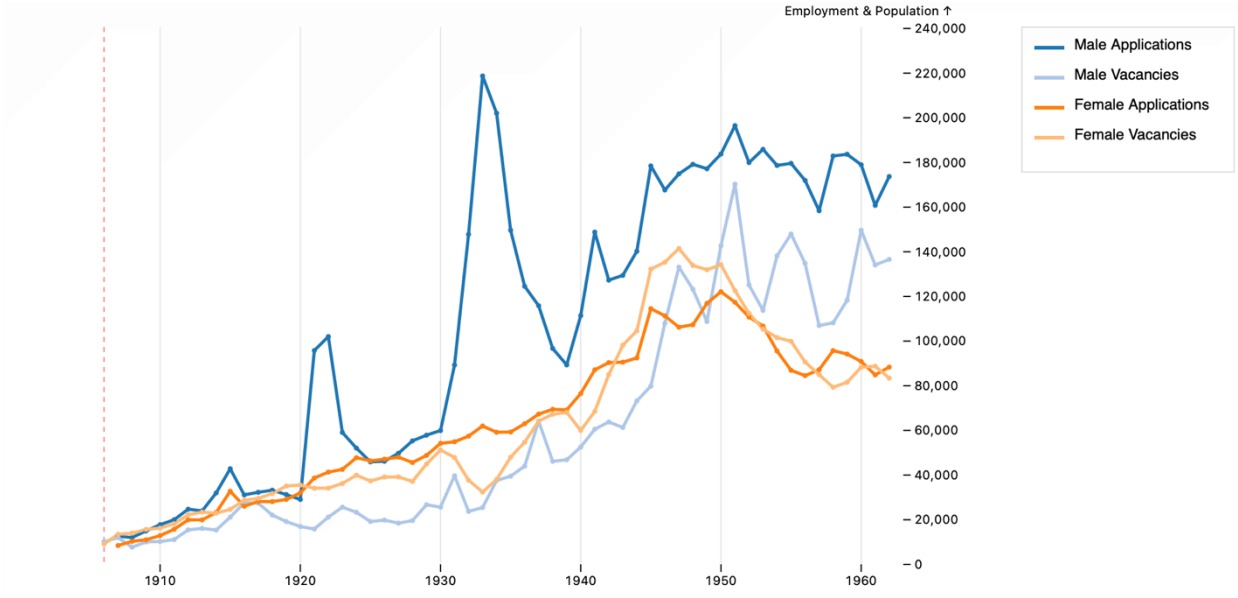
Source: JADB

When assessing the findings obtained from Dagens Nyheter we need to recognize that this was a newspaper that mainly contained ads related to jobs in Stockholm. To contextualize the newspaper-based measures obtained from JADB we have extracted counts of job applicants and vacancies registered at the Stockholm's Public Employment Agency for the period 1906-1962.[5] These statistics are disaggregated by gender.

[5] These figures are reported in the annual publication *Statistisk årsbok för Stockholm*. See Appendix B for further details.

Figure 3 shows the number of applicants and job postings registered with the Stockholm employment office from 1906 to 1966. The overall trend is similar to the one we see in the JADB data, with a steady increase until the mid-twentieth century. However, there are also some interesting differences. For example, the number of male applicants shows sharp peaks in 1921-1922 and 1932-1935, probably reflecting the recessions in Swedish economy during the 1920s and 1930s. These peaks are not as pronounced in the JADB data. Identifying the causes of these divergences is a subject for future research. For now, all we can do is hypothesize why these two peaks are not pronounced in the JADB data. One possibility is that jobseekers did not advertise as frequently in newspapers as companies with vacancies; we further need to remember that placing an ad in the paper cost money, and jobseekers might simply not afford to place ads. Studying early-twentieth century United States, Bi et al. (2025) have shown that higher advertisement prices reduced the number of ads by job seekers, especially by women.

Figure 3. Open vacancies and applications listed at Stockholm's Employment Agency, 1906-1962



Source: See Appendix B.

2.5 Caveats and Limitations

The first caveat, as mentioned above, is that our JADB-prototype collects the number of page-hits, not actual ads. As far as we have been able to manually verify these hits, we are getting

exclusively ads; but we cannot rule out the possibility that some non-advertisement texts were included. It is also very likely that advertisements used other occupational titles in addition to what we found in HISCO. Nevertheless, it is very likely that high-volume gets undercounted with this approach due to a combination of many ads appearing on a single page and ads of multiple positions on a single go.

With this level of abstraction, it is therefore not possible to say how many ads nor how many vacancies were posted. Further, in the current version of the database we cannot distinguish between instances when companies were looking for workers and vice versa. Though our hermeneutic readings indicate that the former far outweighs the latter. These timelines should therefore be interpreted as proxies – signals rather than direct counts for studying the underlying labour market.

There are significant barriers to performing robust data quality assurance: firstly, is the volume of the data – we have identified 69,782 page-hits for the 'Messenger' occupations (HISCO code 37040) alone – and in total 2,054,145 page-hits across all occupations for the period 1870-1970. Secondly, we do not get access to the actual pages – only the number of page hits and, at most, links to the first 20 pages.

2.6 Accessing JADB data

To encourage others to use the data, we have published public, interactive dashboards with different subsets of the data for anyone to explore and export. The latest, and most complete, version of this dashboard is available at https://jobadsdata.github.io/jadb-1-timelines/ where the user can select which data to plot, zoom in on specific parts and export the data, plot and static URLs for any of the plots. An alternative approach to the HISCO-based data is available at https://jobadsdata.github.io/HISCOTrend/ where the user can browse a broad subset of pre-processed interactive plots and search for specific titles. A similar site using only the data for the unit group Messenger (HISCO code 37040) is available at https://jobadsdata.github.io/JADB-37040/; For a short set of example titles, see Table 3. We have also published the output from an earlier experiment, where we focused on youth employment – but otherwise identical to JADB -- at https://jobadsdata.github.io/YouthTrend/.

3 Concluding remarks

The Job Ads Database (JADB) prototype is a new resource that offers a unique perspective on the historical labour market in Sweden. By providing a consistent, long-term dataset on Swedish job ads, the JADB prototype enables researchers to explore a wide range of questions about the changing nature of work, the evolution of recruitment practices, and the dynamics of the labour market over the past century. We hope that this paper will encourage other researchers to use the JADB and to contribute to the further development of this valuable resource.

Using this prototype as a foundation, our intention is to expand the database in both width and depth: successively adding more newspapers while simultaneously disaggregating the data into individual job ads. This will enable the mapping of broad national patterns as well as and to explore the unique histories of individual occupations.

References

Arbetsförmedlingen. (n.d.). *100 år av Platsjournalen*. https://arbetsformedlingen.se/download/18.3e623d4f16735f3976e1379/1543412920091/platsjournalen100ar.pdf

Battu, H., Seaman, P., & Zenou, Y. (2011). Job contact networks and the ethnic minorities. *Labour Economics*, *18*(1), 48–56. https://doi.org/10.1016/j.labeco.2010.07.001

Bentolila, S., Michelacci, C., & Suarez, J. (2010). Social contacts and occupational choice. *Economica*, *77*(305), 20–45. https://doi.org/10.1111/j.1468-0335.2008.00717.x

Bi, H., Petrosky-Nadeau, N., Traum, N., & Woodward, G. (2025). Worker and Firm Search in the Labor Market: Evidence from Classified Advertisements. *Federal Reserve Bank of Kansas City Working Paper No. RWP25-07*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5380732

Bramoullé, Y., & Saint-Paul, G. (2010). Social networks and labor market transitions. *Labour Economics*, *17*(1), 188–195. https://doi.org/10.1016/j.labeco.2009.09.011

Calabrese, M., & Van Leeuwen, B. (2023). *Socio-economic characteristics as determinants in the job market: The case of Piedmont in Italy (1867–2005)*. https://mpra.ub.uni-muenchen.de/id/eprint/119299

Calvó-Armengol, A., & Jackson, M. O. (2004). The effects of social networks on employment and inequality. *American Economic Review*, *94*(3), 426–454. https://doi.org/10.1257/0002828041464542

Calvó-Armengol, A., & Zenou, Y. (2005). Job matching, social network and word-of-mouth communication. *Journal of Urban Economics*, *57*(3), 500–522. https://doi.org/10.1016/j.jue.2005.01.001

Deming, D., & Kahn, L. B. (2018). Skill requirements across firms and labor markets: Evidence from job postings for professionals. *Journal of Labor Economics*, *36*(1), 337–369.

Edvinsson, S., & Westberg, A. (2024). *Swedish occupational titles from CEDAR [dataset]* [Data set]. IISH Data Collection. https://datasets.iisg.amsterdam/dataset.xhtml?persistentId=hdl:10622/KNGX6B

Fishback, P. V. (1998). Operations of ”unfettered” labor markets: Exit and voice in American labor markets at the turn of the century. *Journal of Economic Literature*, *36*(2), 722–765.

Gladoić Håkansson, P., Karlsson, T., & La Mela, M. (2023). Running out of time: Using job ads to analyse the demand for messengers in the twentieth century. *Scandinavian Economic History Review*, *71*(3), 299–318.

Haffenden, C., & Sikora, J. (2025). *Doing digital research at KBLab: A practical introduction to using the National Library of Sweden's data lab*. University of Tartu Library. https://doi.org/10.58009/AERE-PERENNIUS0171

Håkansson, P., & Tovatt, C. (2017). Networks and labor market entry–a historical perspective. *Labor History*, *58*(1), 67–90.

Holzer, H. J. (1988). Search method use by unemployed youth. *Journal of Labor Economics*, *6*(1), 1–20. https://doi.org/10.3386/w1859

Karlsson, T. (2019). Databasen Svenska dagstidningar–mycket text, mindre kontext. *Historisk Tidskrift*, *139*(2). https://www.historisktidskrift.se/index.php/june20/article/download/291/245

Kleinert, J., & Moelzer, W. (2024). *Emergence of the Austrian labor market*. https://static.uni-graz.at/fileadmin/_files/_project_sites/_historical-job-ads/Emergence_Austrian_labor_market.pdf

Kungliga biblioteket. (n.d.). *Svenska tidningar*. Svenska tidningar. Retrieved 15 October 2025, from https://tidningar.kb.se/

Leontaridi, M. (1998). Segmented labour markets: Theory and evidence. *Journal of Economic Surveys*, *12*(1), 103–109.

Licht, W. (1992). *Getting work: Philadelphia, 1840-1950*. University of Pennsylvania Press.

Ljungberg, J. (2025). European consumer price indices since 1870. *Cliometrica*, *19*, 29–80.

Piore, M. J. (1979). Introduction. In M. J. Piore (Ed.), *Unemployment and inflation: Institutionalist and structuralist views*. White Plains, N. Y.

Ros, R., Van Erp, M., Rijpma, A., & Zijdeman, R. (2020). Mining wages in nineteenth-century job advertisements: The application of language resources and language technology to study economic and social inequality. *Proceedings of the Workshop about Language Resources for the SSH Cloud*, 27–32.

Rosenbloom, J. L. (1994). Looking for work, searching for workers: US labor markets after the Civil War. *Social Science History*, *18*(3), 377–403.

Schulz, W., Maas, I., & Van Leeuwen, M. H. (2014). Employer's choice–Selection through job advertisements in the nineteenth and twentieth centuries. *Research in Social Stratification and Mobility*, *36*, 49–68.

SwedPop. (2022, July 4). *SwedPop Documentation: HISCO codes and description*. https://swedpop.se/wp-content/uploads/2022/07/HISCO-codes-and-description-for-IDS-v2.pdf

van Leeuwen, M., Maas, I., Miles, A., Edvinsson, S., Karlsson, J., Jarnaes-Erikstad, M., Pelissier, J. P., Rébaudo, D., de Sève, M., & van de Putte, B. (2002). *HISCO. Historical international standard classification of occupations*. Leuven University Press. https://hal.science/hal-02827799/

Venglarova, K., Adam, R., Mölzer, W., Balasubramanian, S., Kleinert, J., Füllsack, M., & Vogeler, G. (2024). *Who Advertises in Newspapers? Data Criticism in Mining Historical Job Ads*.

Appendix A. Example Query

```
curl -X 'GET' \
'https://data.kb.se/search/?q=%22<MARKER1>%20<OCCUPATION>%22~5%20OR%20%22<MAR
KER2>%20<OCCUPATION>%22~5&searchGranularity=part&isPartOf=Dagens%20nyheter&li
mit=1&offset=0&number_of_fragments=11' \
-H 'accept: application/json'
```

Where `<OCCUPATION>` represents a spelling of an occupational title and `<MARKER1>` and `<MARKER2>` represent different markers. The above URL is equivalent of searching for: `"<OCCUPATION> <MARKER1>"~5 OR "<OCCUPATION> <MARKER2>"~5` That is, it will look for all occurrences of the specific spelling of the occupational title that appeared within five words from either `MARKER1` and/or `MARKER2` across all digitized issued of *Dagens Nyheter*.[1] The response is a JSON file that lists the first page fitting this criterion and a lot of metadata related to the search, including the temporal distribution in the form: number of page hits per year.

Replicating the above process is a fairly simple programmatical challenge, but in the name of transparency and in the hope that others should not have to reinvent the wheel, so to speak, we have cleaned up our specialized script that dealt with our data exclusively to create a general solution: a simple CLI tool we dubbed `kb-trend` – available on PyPI[2]. The underlying code is published under a CC BY NC license on GitHub[3], where we also have a Single Page App version that can be used for one-off queries to retrieve trends as a .csv file.

The resulting JADB dataset records the number of pages with job ads for each occupation title and year. Since all occupation titles in HISCO are classified by gender (male, female or both) we can also disaggregate the job ads by gender.

---

[1] Note the literal name for the newspaper is “Dagens Nyheter”, commonly called simply “DN”, but in the newspaper corpus it is referred to as “Dagens nyheter”.

[2] https://pypi.org/project/kb-trend/

[3] https://github.com/DigitalHistory-Lund/kb-trend

Appendix B. Data and sources for Figure 3

Year	Male		Female		Source information	
	Applications	Vacancies	Applications	Vacancies	*Statistisk årsbok för Stockholms stad* (year)	Page(s)
1906		10,159		9,133	1909	406
1907	12,639	119,85	8,451	13,494	1909	406
1908	12,023	77,19	10,350	14,000	1909	406
1909	14,993	10,077	11,036	15,662	1909	406
1910	17,791	10,248	12,897	16,251	1913	306
1911	20,000	11,120	15,733	18,225	1913	306
1912	24,724	15,467	19,959	22,238	1913	306
1913	23,968	16,115	19,935	23,430	1916	287
1914	32,017	15,389	23,241	22,916	1916	287
1915	42,832	21,163	32,772	24,649	1916	287
1916	31,126	27,541	26,061	28,739	1920	226
1917	32,304	27,427	28,173	29,569	1920	226
1918	33,226	22,088	28,128	31,718	1920	226
1919	31,275	19,193	29,071	35,104	1920	226
1920	29,066	16,970	32,019	35,522	1925	115
1921	95,790	15,859	38,698	34,114	1925	115
1922	102,023	21,176	41,337	34,149	1925	115
1923	59,075	25,640	42,561	36,251	1925	115
1924	52,075	23,320	47,795	39,925	1925	115
1925	45,922	19,186	46,365	37,368	1929	144
1926	46,209	19,841	47,128	39,127	1929	144
1927	49,754	18,475	48,025	39,179	1929	144
1928	55,332	19,628	45,615	37,162	1929	144
1929	57,882	26,738	48,809	44,961	1934	128
1930	59,942	25,622	54,212	51,219	1934	128
1931	89,306	39,704	54,908	47,907	1934	128
1932	147,807	23,751	57,449	37,702	1934	128
1933	218,739	25,446	61,981	32,416	1934	128
1934	202,072	37,587	59,182	38,041	1939	209
1935	149,654	39,503	59,281	48,037	1939	209
1936	124,552	44,009	62,987	54,727	1939	209

Year	Male		Female		Source information	
	Applications	Vacancies	Applications	Vacancies	*Statistisk årsbok för Stockholms stad* (year)	Page(s)
1937	115,828	64,138	67,307	64,187	1939	209
1938	96,693	46,171	69,443	67,235	1939	209
1939	89,344	46,852	69,185	68,069	1944	213
1940	111,423	52,554	76,508	59,983	1944	213
1941	148,836	60,522	87,136	68,483	1944	213
1942	127,302	63,790	90,366	84,990	1944	213
1943	129,435	61,300	90,558	98,161	1944	213
1944	140,210	73,223	92,454	104,568	1948	205
1945	178,460	79,828	114,537	132,247	1948	205
1946	167,655	107,967	111,322	135,318	1948	205
1947	174,881	133,125	106,231	141,453	1948	205
1948	179,171	123,189	107,329	133,865	1953	207
1949	177,184	108,775	116,901	131,955	1953	207
1950	183,764	142,714	122,095	134,237	1953	207
1951	196,503	170,305	117,401	122,633	1953	207
1952	179,947	125,170	110,757	112,366	1953	207
1953	185,877	113,742	106,747	105,398	1955	199
1954	178,665	138,154	95,584	101,573	1956	201
1955	179,602	147,989	86,932	99,922	1956	201
1956	171,949	134,805	84,532	90,647	1958	218
1957	158,369	106,966	87,179	84,998	1958	218
1958	182,910	108,199	95,734	79,321	1960	230
1959	183,657	118,310	94,210	81,565	1960	230
1960	178,991	149,673	90,884	88,289	1962	257-258
1961	160,765	134,101	84,869	88,745	1962	257-258
1962	173,694	136,579	88,355	83,440	1963	257-258

Lund Papers in Economic History ISRN
LUSADG-SAEH-P--26/264--SE+19

Orders of printed single back issues (no. 1-65)
Department of Economic History, Lund University
Postal address: P.O. Box 7083, S-220 07 Lund, Sweden
Telephone: +46 46 2227475
Telefax: +46 46 131585

Full-text electronic issues (no. 58, 60, 61, 66--)
www.ekh.lu.se

Lund Papers in Economic History are published by the Department of Economic History, Lund University, Sweden. This series replaces the former series under the title *Meddelande från ekonomisk-historiska institutionen, Lunds universitet*. The change of name reflects the orientation of the series towards an international readership. The series is multilingual, but the majority of the working papers appear in English.

Lund Papers in Economic History are published as occasion arises, not at fixed intervals. Printed issues are distributed to libraries. From 1999 and onwards, full-text electronic issues are also available on *www.ekh.lu.se.* Those who would be interested in receiving information by email on new issues of **Lund Papers in Economic History** are requested to send an email message to *Lund.Papers@ekh.lu.se.*

LUND
UNIVERSITY