FOR RESEARCHERS

ZARAH DB is a collection of meta data including fairly elaborate research data on, and digital reproductions of, archival, published and interview sources on the broad theme of *Women's Labour Activism in Eastern Europe and Internationally from the Age of Empires to the 1990s*. The database is associated with the research project ZARAH, ERC Advanced Grant (Grant agreement No. 833691, 2020-2025). ZARAH explores the history of women's labour activism and organizing to improve the labour conditions and life circumstances of lower and working-class women and their communities. It discusses local, nation-wide, cross-border and international contexts, networks and organizations, and incorporates a rich variety of activisms, ranging from grass-root initiatives to social and political action within public and private institutions, and from anarchist radicalism and working women's involvement in co-operatives and trade unions to cross-class women's organizing.

Through a highly curated set of keywords, abstracts and records on institutions, events, people and places, and an intuitive search interface, ZARAH DB presents and interlinks materials whose multilingual originals are otherwise held in dozens of public and private institutions and collections across and beyond the region studied. We are particularly grateful for the cooperation of the institutions holding original materials and the support of their staff, and carefully document the original locations of all materials included in this database.

Within ZARAH DB, English-language research data complement and supplement machine-recognized content written in the various languages and related to the countries represented in the ZARAH component studies, including Austria, Bulgaria, the Czech lands, Hungary, the Polish lands, Romania, Turkey, and the Yugoslavian lands; these documents are accompanied by key sources of a transregional and international character.

This collection is meant to be of service to the scholarly community as well as to anyone interested in women's labour activism. We hope that both specialists and the broader public will make use of ZARAH DB to gain a sense of the complexity and variety of materials and information held by archives and institutions across the ZARAH region and internationally, on the under-explored topic of women's labour activism in Central and Eastern Europe and beyond. We also hope that ZARAH DB will contribute to the conversation on the past and present of such activism. In addition, ZARAH DB (with its public functions and those not visible to the public) operates as a collaborative research tool serving the needs and interests of the ZARAH project.

ZARAH DB is innovative because it includes documents about underexplored and marginalized persons, events and processes, uses categories and keywords shaped by cutting-edge research on gender and labour and tailored to reflect the history of the region studied, and finally, because it offers insight into the research and collaboration process of a large team of researchers studying a complex subject and documents in many languages.

ZARAH DB collects a wide breadth of documents relevant for the history of women's labour activism. However, we would like to emphasize that the documents and information included in ZARAH DB are per force, selective. The database builds closely on and mirrors our collective research process, while aiming to include selected key documents that also lead beyond the ZARAH project.

In what follows you can find: specifics about the scope and forms of access to documents and meta data, including our fairly elaborate research data, in our growing publicly-accessible collection (WHAT WE HAVE); details about the types of keywords making up the English-language metadata that constitutes an essential part of each entry and, together with the document abstracts, the backbone of our Advanced Search system (HOW OUR DATA IS STRUCTURED); and information about how to search and filter data in ZARAH DB (HOW TO USE).

WHAT WE HAVE

Scope

ZARAH DB contains materials collected by ZARAH project team members, which have been used to elaborate their component studies or their contributions to one of the collaborative volumes of the project. The ZARAH researchers create bibliographic entries and English-language annotations, and, when possible, make public digital reproductions of all original materials they use. The ZARAH researchers also aim to include selected key documents leading beyond their immediate work with sources.

The database will continuously acquire entries until the end of the project, in the year 2025.

Those using the database may expect to find a growing number of documents in ZARAH DB, as well as attached information drawn from the following types of sources:

- Trade union and labour organizations' sources at local, national and international levels
- Women's movements' and organizations' materials of a local, national and international outlook
- Archival material of non-trade union organizations where women pursued labour politics
- Other institutional records and documents, including police records, legal records, records of the labour inspectorate, party and state documents
- National and local press, including women's and labour movement journals and factory newspapers
- Workplace-based and local archival documentation
- Visual material contained in these types of documents or discovered elsewhere
- Ego documents
- Interviews with labour activists

Users can expect to see entries referring to originals held in repositories from the countries/lands specified above, covering the region dealt with in ZARAH component studies and collaborative volumes. In addition, ZARAH DB contains and will continue to add entries referring to select original materials held in archives specializing in the history of women, labour and social movements, or preserving the records of international organizations, from countries such as Germany, the Netherlands, Russia, Switzerland, the United Kingdom and the United States of America.

Types of data and levels of accessibility

All items present in the database have, at the least, a detailed bibliographic reference (full citation), keywords and a short abstract attached. The variation in the amount of information visible about each entry is shaped, in part, by the preferences expressed by archival

institutions and interview subjects when granting reproduction permissions, and by the project's commitments to privacy rights and non-exposure of sensitive data – in accordance with the EU's General Data Protection Regulations (GDPR) and other applicable norms.

Specifically, a digital reproduction of a source document is available to the public only if the institution holding the original has granted ZARAH and ZARAH DB explicit written permission to display that reproduction, and if all copyright-related questions have been clarified. In the case of interviews conducted by ZARAH team members, the database may contain an interview transcript, the audio recording of the interview, both items or neither – depending on the choices made by an interviewee concerning anonymity and the storing of their data.

For details about our terms of use and personal information protection, please consult the dedicated pages: https://public.zarahdb.eu/terms-of-use and https://public.zarahdb.eu/privacy-notice.

The note "PDF – Restricted access" denotes the absence of a publicly-visible digital reproduction of the original document (which might, however, be added at a later stage); a full citation and the meta data including our fairly elaborate research data in the English language concerning that document are still available. If a digital reproduction has already been published by another institution, we link to it or refer to the repository.

If a document is typeset (rather than handwritten), it is processed before being uploaded with the help of optical character recognition (OCR-ing) software. The text in the document becomes searchable within ZARAH DB.

HOW OUR DATA IS STRUCTURED

Document entries

The database contains searchable entries. A specific entry consists of bibliographic information, structured (pre-set) as well as individual (free) keywords, and an English-language abstract describing an original document, and (if the holding institution has granted permission and/or the publication is not under copyright) a digital reproduction of the document. (We continuously create new entries and acquire additional permissions to display digital document reproductions.)

Keywords in ZARAH DB can be divided into: "Special Record" keywords, structured (pre-set) keywords and "free" keywords set in addition to the structured keywords, which direct researchers with specific thematic interests to specific entries.

"Special Records": People, Places, Events, Organizations

In ZARAH DB, Special Records are keywords referring to key Persons, Places, Organizations and Chronology (Events). They are different from the pre-set and free keywords because they often have additional descriptive information attached to them; this additional information is styled in the manner of Authority Records created by libraries for their catalogs. Special records on organizations and people give unified information on organizations and individuals mentioned in the document entries. Key events and place names are also recorded systematically. This unified system allows researchers to search ZARAH DB for documents and additional information about a person who might have used

different names during her (his) lifetime or a place that might have had various names in different languages over time, by entering only one of the known names. The special recording of events allows for systematic searching among and chronological ordering of the key events reflected in the documents included in ZARAH DB.

The four types of Special Records within ZARAH DB contain different types of information and may display specific formatting:

▪ *People* records

People records contain biographical information about the person and may collect references for further reading about that person. Importantly, these types of records collect, where possible, known alternative names and pseudonyms of persons appearing in a document.

People records help uncover and spotlight the contributions of a wide range of labour activists. As they become attached to entries, they help display links between individual labour activists and various key events or organizations, and the links of various other actors (journalists, politicians etc.) to women's labour activists, and the activists input to representing or shaping the agendas and repertoires of women's labour activists.

▪ *Organization* records

Organization records contain unified information about an organization or part of an organization referenced in one or more documents.

When viewing the sliding information card about organizations (see below), it is possible to see additional information about the organization's usual acronym, organizational form, the scale of its operations, and whether its membership was gendered and in what way.

These pre-set categories attached to Organization records contribute to making visible the variety of structures and sheer diversity of organizations in which women's labour activism took place, or that were addressed or challenged by activists.

Please note that the original (language) name of an organization may not always be displayed (i.e. when we have not been able to establish it beyond doubt). Also, the country or territory name provided in parentheses after the name of an organization is a simplified version of the country's or territory's name at the time when the document to which the Organization record is attached was created.

▪ *Event* records

Event records show, at a glance, English language information about events (e.g., strikes, committee meetings, trade unionist functionary-interviews with factory workers, etc.) at the core of the document, or being referenced in the document. Together, these records constitute an evolving chronology of events that were shaping women's labour activism in our region and beyond, within the limitations arising from the specific focuses of our individual and collective research, which are mirrored in the documents included in ZARAH DB.

The country or territory name provided at the end of an event's short description follows the style guide used for the Places records.

- *Places* records

Keywords associated to the Places category are styled to name both the sovereign entities (and territories, if applicable), cities or villages mentioned in the document, and the current name of the sovereign entity of that location. The present-day name is given in a short version and in square brackets, e.g, [Turkey].

This notation system has been devised to adequately describe the complex history of place names in the region studied by ZARAH. The system allows users to find documents in ZARAH DB that refer to a place that might have had various names in different languages and/or over time, by entering only one of these names.

Pre-set keywords in the document entries

The ZARAH project's core conceptual framework is reflected in the pre-set keywords (controlled vocabulary), which are used to describe (annotate) a document. The keywords reflect the conceptual framing of ZARAH contained in the project proposal as submitted to and accepted by the ERC, as well as some further development and refinement in the initial phase of the project.

When annotating a document, we use keywords that refer to the following dimensions of analysis:

- Repertoire (of action), with the sub-categories:
    Action (e.g., classical – strike, alternative – boycott, non-activism/choice not to act)
    Scale of repertoire action (e.g., local, national, transnational)
    Format of participation (e.g., women-only, trade union, cooperative)
    Communication and knowledge production (e.g., reporting, public speaking)

The pre-set keywords grouped under "Repertoire (of action)" are meant to help capture and make visible online the full array of women's labour activism in (and beyond) ZARAH's region of focus, the diversity of scales and formats in which it occurred, and the variety of communication genres employed in (and as) activism. The sub-categories bring to light, besides the classical forms of labour activism, the non-trade union, informal, sporadic, issue-based character of women's participation, and actions that have been often misconstrued by trade unionists and historians alike as "insignificant", "spontaneous", as "lacking class consciousness" or born out of a "reluctance to organize". Importantly, they foster recognition of the historical specificities of women's labour activism in ZARAH's region of focus. For instance, we have included pre-set keywords such as "activism within the state" (e.g., for reformist actions of women trade unionists active in state-socialist regimes) or the sub-category "non-activism/choice not to act" (a conceptualization which recognizes agency and rationality in some women labour activists' inaction in certain moments or contexts). Also, the keywords included under "repertoires (of action)" enable ZARAH researchers to spotlight the gendered character of labour organizing, as well as the importance and frequency of mixed-sex organizing in working class and other dominated communities.

- Agendas (e.g., income and wage justice – equal pay for equal work, social rights, welfare politics and the politics of the double day – demands for/related to maternity leave, family allowances, other social programs);

The pre-set keywords in the Agendas category refer to the types of claims made by women labour activists. They build on existing historical and social research on social movements on gendered activism. The keywords from each sub-category are meant to go beyond capturing the typical demands of male trade unionists, striving to capture both demands that dealt with workplace injustice and demands related to the everyday lives of communities. They reflect a conceptualization of labour that includes both paid and unpaid work (including housework) and an awareness of the burden of care historically placed on women.

With the help of these, as well as additional keywords not displayed publicly, ZARAH also aims to generate knowledge about long-term trends, cycles, and waves in the history of women's labour activism in our region and beyond.

- Effects of activism (e.g., full goal achievement [what was requested is exactly what occurred], adverse effects);

The keywords in this sub-category use the operationalization for "goal achievement" applied by Mary C. Joyce in her MA thesis. See Mary C. Joyce, "Activism Success: A Concept Explication" (MA Thesis, Seattle, USA, University of Washington, 2014), 69, http://hdl.handle.net/1773/26455, accessed 13 June 2021. In addition to Joyce's original conceptualization, the keyword "adverse effects" was added to this sub-category, so as to strengthen the fit of Joyce's operationalization for the type of material analyzed within the ZARAH project.

- Historical context (e.g., economic crisis, post imperialism, state socialism);

The pre-set keywords in this category are meant to provide a shorthand way of capturing important large-scale historical phenomena and changes in the period represented in the ZARAH project and in ZARAH DB. These keywords reflect the premise that our region of focus is part of a world region that was characterized by unequal development and differentiation. These broad keywords are meant to aid in understanding how ZARAH DB materials related to region-specific traces of large historical processes.

- Labour conditions (e.g., seasonal work, rationalization, discrimination);

The keywords in this sub-category refer to labour conditions frequently encountered by women workers since the 19th century, particularly in the project's region of focus. They help reflect the historical prevalence of women's paid but often informalized and highly exploitative labour, and the frequency of phenomena such as intense labour migration, or intensity of processes like rationalization and deskilling.
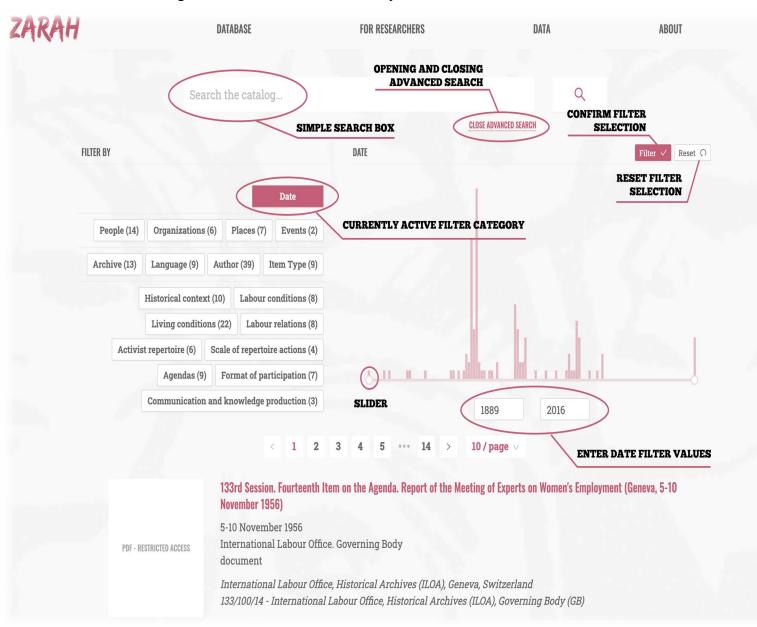
- Living conditions (e.g., housing, authorities – police hostility, intra-family and intra-community issues – gendered violence);

The keywords in this sub-category refer to various types of problems linked to housing, social infrastructure (education, healthcare, childcare, availability of basic goods), technical infrastructure (transport, sanitation, ecology), relations with authorities (police, other forms of authority creating insecure environments) and intra-family and intra-community issues (burdensome, unpaid family work, gendered violence, generally insecure social relationships). The existence of these pre-set keywords draws attention to the variety of living conditions that determined or shaped women's labour activism, in and beyond formal workplaces.

- Labour relations (e.g., reciprocal labour-household(s) – kin non-producers, market commodified work – wage earners – time-rate);

The keywords in this sub-category use the "Taxonomy of Labour Relations" developed by the *Global Collaboratory on the History of Labour Relations, 1500-2000* project hosted by the International Institute of Social History, led by Karin Hofmeester. For details on the development of this taxonomy, see Hofmeester, Karin, Jan Lucassen, Leo Lucassen, Rombert Stapel, and Richard Zijdeman. "The Global Collaboratory on the History of Labour Relations, 1500–2000: Background, Set-Up, Taxonomy, and Applications." International Institute of Social History, 2016. Collaboratories in Social and Economic History. http://hdl.handle.net/10622/4OGRAD.

We have incorporated this taxonomy due to its comprehensiveness and capacity to capture a variety of paid and unpaid labour arrangements, and to connect our findings on labour relations involving women in a particular world region to this formidable and important collection of global research data on the history of labour relations.

HOW TO USE

Simple Searching

You can search the ZARAH DB collection by performing a Simple Search - typing one or several keywords of interest in the search engine on the ZARAH DB front page (landing page). (The Simple Search box can also to be found at the top of the "Database" website section).

When performing a simple keyword search, the engine searches through all the text in the OCR-ed document reproductions in the database (if they are visible to the public), all the keywords attached to entries and any other text (abstracts, citations, detailed information about People, Places, Organizations or Events).

In this context, please note that ZARAH DB is a multilingual database, in which documents are written in various languages present in the ZARAH region, while all metadata is in the English language. Consequently, simple searches for terms in languages other than English will yield fewer results (but by no means none) compared to searches for keywords in the English language.

Advanced Searching

Advanced searching is available by building on or independently from simple searching.

You can sort through the results of a simple search by using the Advanced Search feature. Click on the "Show Advanced Search" link under the Simple Search engine box to access this feature.

You may also use the Advanced Search as an alternative, more structured, type of sorting through documents and metadata, compared to the Simple Search.

In Advanced Search you can filter data according to the key research categories (detailed in the "How our data is structured" section above). You can also filter materials according to basic bibliographic information: author name, item type (newspaper article, document, journal article, interview etc.), archive (institution holding original document), language (of the original document).

Besides filtering, in Advanced Search you can also access additional information about People, Places, Events and Organizations. Please click on the [+] symbol appearing next to keywords from these four categories; an "information card" (record) will slide out from the right-hand side of the screen.

For the Date filter, please use the slider or the open fields to select, or type in a desired time interval. Use the Filter or Reset buttons (on the top, right hand side, next to the timeline) to set your time interval or to reset it.

When performing Advanced Searches, you can add or remove various filters. Each additional filter further narrows down the results of your initial search. To widen your search results again, click on the X mark to remove a filter. Different combinations of filters will, of course, produce different search results.

Viewing document entries

Results from simple or advanced searches are displayed below the search engine box, in a Search Results List. You can choose to view a document in detail by clicking on its thumbnail preview in the list of results.

When viewing the full entry of a specific document, you can click on the various keywords attached to it to display other documents that have the same keywords attached.

To access any additional information attached to a Special Record keyword, while viewing a specific document, please click on the keywords in the Persons, Places, Organizations or Events category. An "information card" (record) containing further information will slide out from the right-hand side of the screen. This sliding information card system allows users of the public ZARAH DB website to quickly view additional relevant information produced by ZARAH project team members.

When clicking on the keywords attached to the categories People, Places, Events or Organizations a "special record" containing additional information about a Person, Place etc. slides out from the right-hand side of the page, for quick consultation.

Notes on translations, spelling and transliteration

In the Language category of a document entry, the first language(s) mentioned is the language(s) in which the document was written or in which the interview was conducted. The second (or last) language mentioned refers to the language of the metadata and of all translations provided. (This second entry is without exception, English).

Translations of organizations' or institutions' names are given in square brackets, always in English, e.g., *Gewerkschaft der Textil-, Bekleidungs- und Lederarbeiter [Textile, Garment and Leather Workers' Union]*.

ZARAH DB uses US American English spelling for its English-language content. Exceptionally, we use the "labour" rather than "labor" spelling.

Citations and bibliographic information in ZARAH DB are formatted in accordance with the guidelines in the Chicago Manual of Style, 17th edition.

Transliteration from the Cyrillic to the Latin alphabet is done using the simplified Library of Congress transliteration system.

We are interested in your feedback about the content of ZARAH DB and its performance. We welcome your questions about this database. Please email your comments and questions to zarahdb@ceu.edu.

* * *

This essay was written by Alexandra Ghit.

It contains extensive reference to the ZARAH research proposal as accepted by the ERC, and the follow-up work related to ZARAH DB, for which Alexandra carried the lead-

responsibility, and to which, besides the PI, other members of the ZARAH research team made important contributions.

In the period leading up to the submission of the research proposal to the ERC, besides Alexandra, Selin Cagatay, Veronika Helfert, Ewa Maczynska, Ivelina Masheva, Sandra Prlenda Perkovac and Radka Sustrova have, in various capacities and to various degrees, contributed to the development of the ZARAH project. In the early stages of the project development Silke Neunsinger also made important contributions.

Jozsef Bone has implemented the database-driven web application, which serves as the business logic layer for the public and administration interface, also designed by Jozsef.

Zsofia Veszely has contributed, as Susan Zimmermann's research assistant, to the generation of a number of entries in ZARAH DB.

Vienna, 20 June 2021