

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

Dr. Michael Karabinos & Dr. Christian Olesen

158621016Y: Preservation and Access

21 December 2023

Dutch Labour Conflicts and Politics (1950-2019):

Enriching the Labour Conflicts Database for Future Use Cases

Abstract

This paper will consider the Dutch Labour Conflicts (DLC) dataset created by Sjaak van der Velden, and will focus on the period of 1950-2019. This dataset focuses on strikes in the Netherlands, especially when considering labour conflicts and politics. For accessibility to the mentioned dataset, the application of OpenRefine will be used to both clean the data as well as make the data more general accessible.

Keywords: Dataset, strikes, labour conflicts, politics, OpenRefine

Related dataset and documentation ‘Dutch Labour Conflicts and Politics (1950-2019)’ with URL <https://github.com/Preservation-and-Access-2023/Labour-Conflicts> in a ‘Github’ repository.

Dutch Labour Conflicts and Politics (1950-2019): Enriching the Labour Conflicts Database for Future Use Cases

The Dutch Labour Conflicts (DLC) dataset has information on more than 15,000 labour conflicts which occurred in the Netherlands between 1372 and 2019. The DLC dataset is hosted by the International Institute of Social History (IISH) (“Search for strikes”). IISH does not provide a critical introduction on the dataset, so this will be given in the first chapter. In the second chapter, a potential use case of the data for political scientists will be suggested. The third chapter will describe the steps required in cleaning the data and linking it to further data on Dutch politics in order to create the Dutch Labour Conflicts and Politics (1950-2019) (DLCP) dataset. In the fourth chapter, opportunities and pitfalls of the DLCP are discussed. Finally, the fifth chapter contains the conclusions drawn from working with this dataset.

1. Introduction to the Data

The first version of the DLC dataset was compiled by Sjaak van der Velden as part of his research for his 2000 doctoral dissertation *Stakingen in Nederland [Strikes in the Netherlands]*. The scope of this early version of the dataset comprised all labour strikes taking place in the Netherlands between 1830 and 1995. For the purpose of the dataset, Van der Velden defined a strike as ‘an interruption of paid employment by two or more people with the intention of returning to work at a later date’ [‘een onderbreking van het werken in loondienst door twee of meer personen met de bedoeling om in een later stadium weer aan het werk te gaan’] (317). Only strikes occurring on Dutch territory were included, with the exception of strikes which took place in some sectors of shipping and off-shore industry, inclusion depended on the site of business for the companies involved.

1.1. Data Provenance

Van der Velden began maintaining a ‘clippings archive’ [‘knipselarchief’] on strikes in 1972

which, after digitisation in 1990, grew into the dataset attached to his dissertation (23). This dataset consisted of 14,724 entries on strikes, each having 38 associated records (expressed in a columns graph). Apart from data on each strike, these entries include a record for references and a record for memos, with the option to add a strike description and/or continuations from other text records which had been restricted due to a character limit. In a separate dataset, Van der Velden collected information on 373 labour actions which did not fit his definition of a strike. Instances when workers expressed their discontent through demonstrations, sabotage actions, and assaults on managers. This dataset constitutes ‘a rather arbitrary and in any case incomplete selection out of the actual number’ [‘een tamelijk willekeurige en in ieder geval onvolledige greep uit het werkelijk aantal’] of such actions (Velden 316). Van der Velden and IISH have incorporated both of these datasets in the DLC.

The scope of the data included in this project – Dutch labour conflicts occurring between 1950 and 2019 – covers three distinct time periods with respect to data collection. For the period from 1951-1981, data on strikes published by the Central Bureau for Statistics (Centraal Bureau voor de Statistiek (CBS)) was aggregated by Gerrit van Kooten as part of his 1988 dissertation, *Stakingen en stakers* [*Strikes and strikers*]. This dataset was later entirely incorporated into Van der Velden’s database. For the period from 1982-1999, the data was aggregated by Van der Velden based on his clippings archive, academic literature, trade union periodicals, and the data collected by CBS. Data on the years from 1996-1999 were added by Van der Velden after the completion of his dissertation research but before its publication (318). Between 2000 and 2019, the data was gathered by IISH researchers and staff. In 2015 IISH launched the HUB Global Labour Conflicts Dataverse – a collection of labour conflict databases – to which the DLC was added (“HUB”).

1.2. Data Biases

A bias arises within the publicly available DLC as a result of the limited definition of ‘labour

conflict' within the data. Jens Aurich et al., writing on the 'HUB Global Labour Conflicts' Dataverse (of which DLC is both a subset and the foundational dataset), referred to the early development of the DLC through Van der Velden's recordings of official strikes, stating that this 'has the effect of historicising strikes by (mostly) male European unionised industrial wage labourers, this being merely one type of collective labour action within a specific region and temporary context' (171). As a result, unionised strikes became the standard for thinking about workers' resistance throughout history, effectively excluding other forms of struggle, such as walkouts, sabotages and boycotts. These alternative and often more subtle forms of resistance are largely those favoured by marginalised groups such as women and people of colour, as exemplified by the actions of Amazon warehouse workers who had, until recently, been denied the right to unionise (Alimahomed-Wilson and Reese 1). Susan Leigh Star, writing on the problem of standardisation, suggested that standards act to reinforce already existing power imbalances; not only through the exclusion of individuals or groups who cannot be defined within the prescribed boundaries, but in further highlighting the story of those who fit within those boundaries. She wrote, 'Power is about whose metaphor brings together worlds and holds them there' (102). In the context of the DLC, the metaphor for labour resistance becomes the unionised worker. In response to what IISH themselves have identified as the problem of foregrounding strikes in the dataset over other forms of labour conflict, Aurich et al. reported that a series of conferences had been organised and discussion among the contributing scholars remains ongoing (169).

Furthermore, in asking questions about what constitutes a labour conflict, it is important to consider what is meant by labour. Second wave feminists posed this question in the 'Wages for Housework Movement' (figure 1) and through a series of Global Women's Strikes (Federici 1), while the argument for 'reparations' has its foundation in acknowledging

the labour of enslaved people.¹ In response to this problem, HUB project leaders developed a taxonomy of labour relations in collaboration with international scholars, made available for public download as a .csv file on the IISH website (figure 2). This taxonomy defines labour under different categories such as type of exchange, goal of production, and labour relations.



Fig. 1. Lane, Bettye. *A Wages for Housework March*. 1977, Schlesinger Library, Radcliffe Institute, New York.

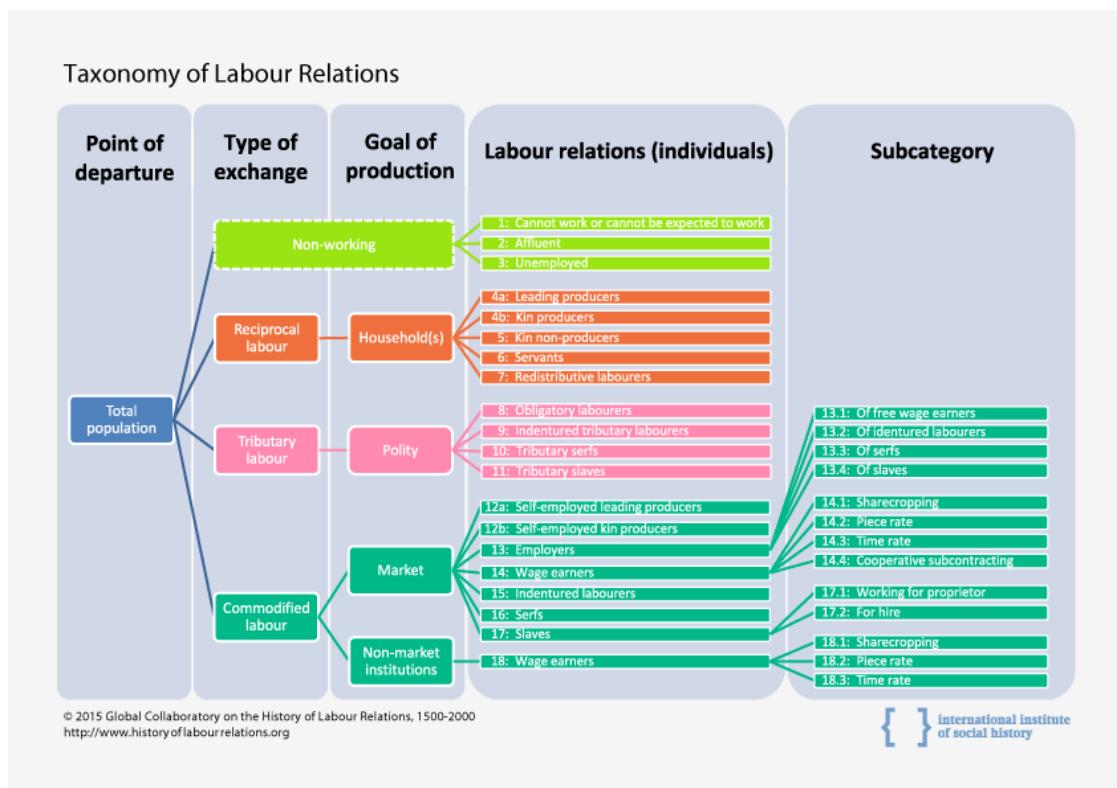


Fig. 2. Taxonomy of labour relations provided by IISH.

¹ Reparations is the term commonly used to define the financial compensation which parties benefitting from the Atlantic Slave Trade – colonising countries, companies and individuals – should be required to pay to the descendants of enslaved people of African heritage (Coates, “Reparations”).

1.3. Data Access

The DLC dataset is initially accessible via the IISH website as a data story, presented as a series of graphs, interactive visualisations (figure 3) and refined searches (figure 4) which focus on factors such as location, duration, outcome and character of strikes.² Instructions on how to interact with the data story are given in both English and Dutch. On the one hand, this makes the dataset highly accessible for potential users and goes some way to addressing what Richard Zijdeman et al. (developers on the CLARIAH (Common Lab Research Infrastructure for the Arts and Humanities) project) have highlighted as a problem of accessibility when working with Linked Open Data and general audiences. They described how CLARIAH's ideal audience of historians are not only 'generally lacking' in programming skills, but are 'literally put off when seeing code or SPARQL code' (1, 10). Thus, by hiding the query itself and presenting the data through clear visualisations, general users have an opportunity to engage with and begin to understand the dataset. On the other hand, what is evident through an initial interaction with the DLC is that the data story predetermines the types of queries that are made. For example, if a user would like to know something about the workers' motivation for striking, it is possible to search either by 'term', which first requires the selection of relevant words – since all descriptions are in Dutch, only Dutch search terms can be used – or by organisation, which brings up a description of strikes, including the reason for striking – for this, it is necessary to already have some knowledge of the relations between an organisation and its employees. For users who are willing and able to work with SPARQL, each visualisation has a link to the SPARQL endpoint where the query which generated that visualisation is shown. This allows users with little experience to experiment with changing such queries one variable at a time.

² The interactive visualisations are arranged as follows: 'Strikes per year', 'Spread of the strikes', 'Which occupations struck the most', 'The top 20 largest strikes since 1900', 'The top 20 companies with the most strikes', 'The outcomes of the strikes', 'Character of the strikes', 'Number of strikes per city since 1900'. (IISH, "Search for strikes").

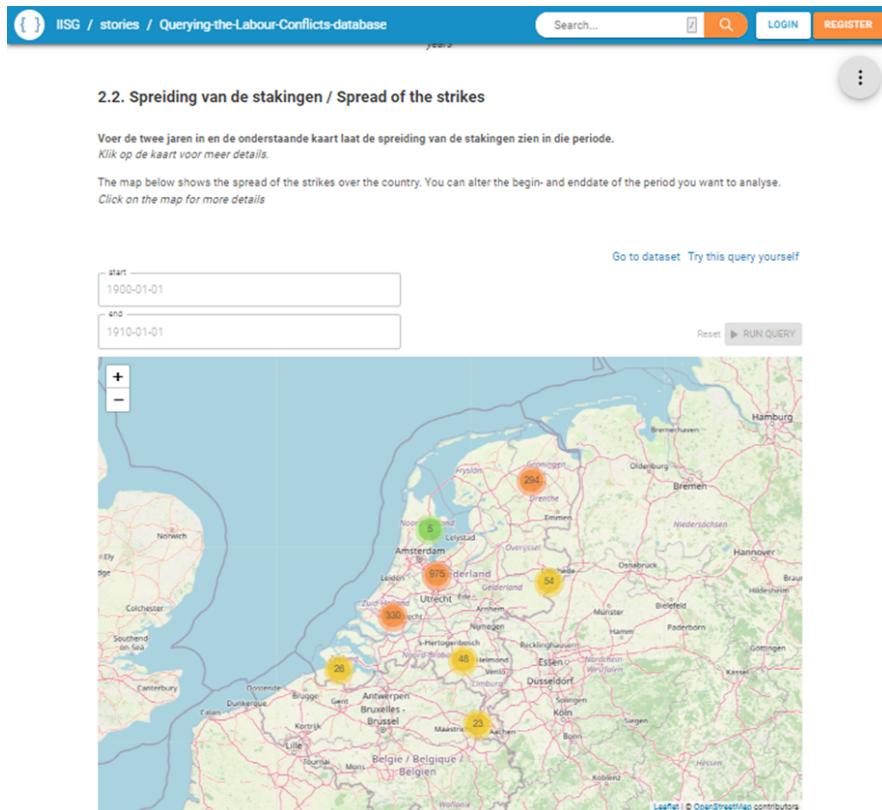


Fig. 3. IISH interactive visualisation 2.2 – part of the Dutch Labour Conflicts data story.

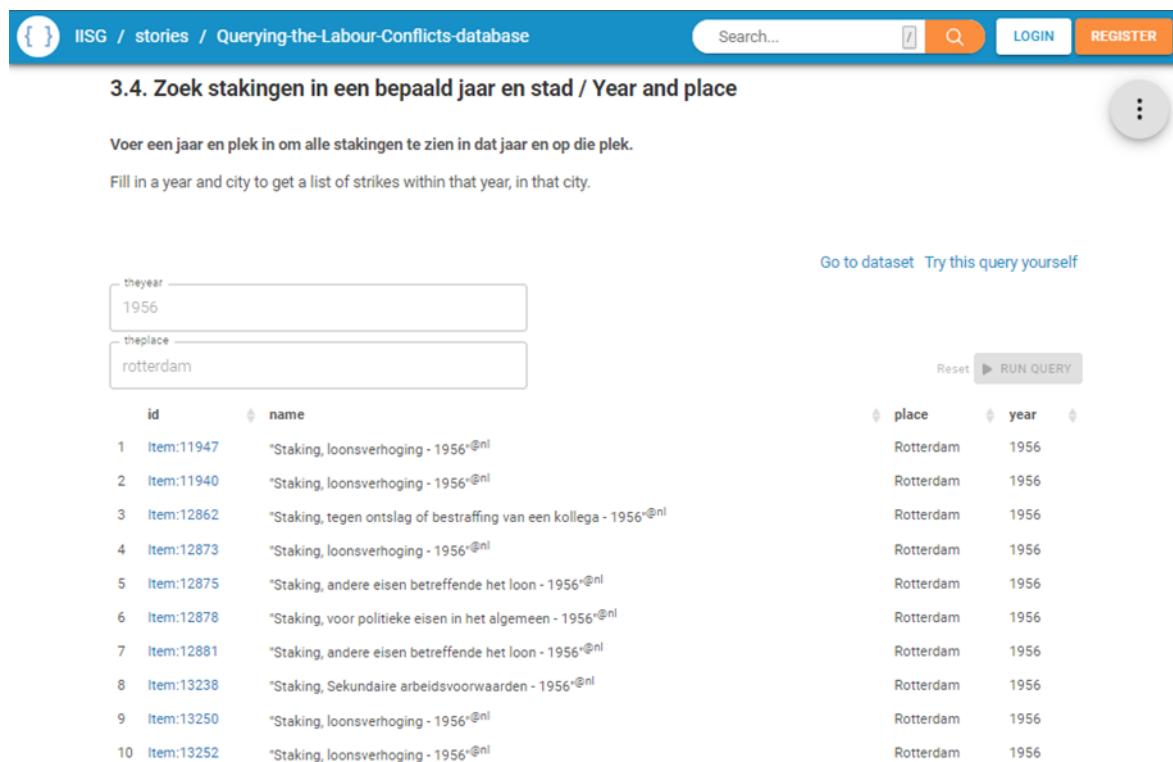


Fig. 4. IISH refined search engine 3.4 – part of the Dutch Labour Conflicts data story.

2. Potential Use Case(s) of the Data

2.1. Problem Statement

The DLC dataset is available both as a graphical representation – iterated in rows and columns – and Linked Data – structured data consisting of subject-predicate-object triples, making it of greater use for semantic queries. In its Linked Data format, the DLC dataset is currently not linked to any objects in external Linked Open Datasets (LOD) such as Wikidata. Jonathan Blaney, in describing the goal of linking data to other datasets, wrote, ‘Analysing huge volumes of data is potentially much more powerful than everyone using their own individual datasets dotted around the web in what are known as information silos. These interoperable datasets are what LOD practitioners are working towards’ (2). He further highlighted the steps needed for creating LOD, following three main principles: using a recognised LOD standard form, referring to an entity in the same way that other people do, and publishing your data openly (2). The unseized opportunity to link the DLC dataset to other datasets means that current research possibilities for the dataset are limited to the research questions posed by the data providers who aggregated the datasets. With further attention and in-depth research, the DLC dataset would benefit from fully implementing Blaney’s three steps for creating LOD. This paper aims to show the strengths and limitations of linking the DLC dataset to other types of data.

2.2. Research Goals

The DLC dataset could be a valuable resource for political scientists and political historians. Historical analysis of labour conflicts in relation to the political landscape may provide new insights into the effects that policies have on broader society, using strikes and other labour conflicts as an indicator of the level of contentment for waged workers. Conversely, it may also provide insight into when and how clearly the voices of wage workers expressed in strikes are heard by the government.

This kind of analysis cannot be undertaken using the DLC dataset on its own, since it does not include any information on the political climate in the Netherlands. The goal of this data paper is to enhance the DLC dataset to enable this kind of analysis. Accordingly, the central research question is formulated as: *To what extent could the Dutch Labour Conflicts database be enriched by linking it to other political datasets?*

By linking DLC with Linked Data about the political landscape, thus creating a Dutch Labour Conflicts and Politics (DLCP) dataset, a myriad of new ways to use the dataset are opened up. A potential question that a political historian could pose is: What, if any, is the correlation between the political leaning of a government and the number and intensity of strikes?

In addressing this question, a problem arises in thinking about how to represent what is meant by political leanings, frequently defined under the terms ‘left’ and ‘right’ and more often than not presented at opposite ends of the political spectrum. In France where the terms originate, the left or ‘liberal’ is seen as favouring change, whilst right or ‘conservative’ is considered to stand for order (Wikipedia contributors, “Left–right”). Yet even when defining where a political party stands on the spectrum (see Wikipedia’s ‘List of political parties in the Netherlands’), politicians sitting in parliament or parties as a whole may change their political stance on different issues. In addition to which, the Dutch political system presents an added difficulty in thinking about the political leanings of a government. Since 1900, no one party has gained enough seats to form a majority in the Dutch parliament, as a result the cabinet has been formed through a coalition of two or more parties.³ In considering the leaning of a Dutch government when enriching the DLC data subset, the decision was made to reference Wikipedia’s ‘List of Cabinets of the Netherlands’, which includes a column labelled ‘political

³ Initially the Dutch Prime Minister’s political stance was suggested as a means to determine the political leaning of a government, but this was later considered to be an inaccurate measurement, owing to the fact that Dutch governments are formed through coalitions. So while a prime minister may remain the same (representing the party with the most seats in parliament), parties in the coalition can change, thus resulting in a different political leaning.

position' under which the cabinets are labelled as anything from 'Left-wing' to 'Right-wing' with 'Centre-left', 'Centrist' and 'Centre-right' occupying the middle of the spectrum.

2.3. Hypothetical Use Case

For more in-depth research, the project DiLiPaD (Digging for Linked Parliamentary Data) could be used instead of the Wikipedia 'List of cabinets' to give a more accurate reading of a government's political leaning. The aim of DiLiPaD is to create a large body of textual data from parliamentary proceedings in three separate countries (UK, The Netherlands, and Canada) in order 'to shed light on developments across different nations, cultures and systems of political representation' (Beelen et al. 3). In one test case, DiLiPaD was used to consider the role of gender in British politics, starting with 677 million words of political speech – drawn from every debate to have taken place in the House of Commons since 1945 – computerised text mining was used to enable a search for specific words in order to determine the political positioning of parties and individuals (Blaxill et al. 412). While the aim of the test case was to consider whether increasing the number of women in parliament directly collated with a greater representation of women's 'issues' and 'interests', the overall approach provides a model for thinking about how parliamentary proceedings in general and the language used during said proceedings might give a clearer view on the political leanings of a government. The analysis began at the simplest level of measuring direct references to women in speech, in the first instance, mention of the word 'women' or 'woman' and in the second instance, tracking a selection of words which were considered to be synonymous with women in a contemporary context (Blaxill 425-6). In addition the frequency of words was deemed to be of importance, allowing for the researchers to determine whether women were mentioned just in passing or if the debate could be said to be 'about women' when the number of mentions exceeded a set threshold (Blaxill 428). The research also considered the gender of the person speaking and which of the three main political parties that person

represented – Labour (centre-left), Conservative (centre right) and Liberal Democrat (centre to centre left).

In reflecting on how DiLiPaD could be used in a similar manner to empirically analyse words spoken in political debate in the Dutch Parliament (Tweede Kamer), the following steps could be taken: Selection of words considered to relate to left or right leaning issues (20 words were chosen in the test case), indication of a threshold for relevance (the test case suggested ten uses of the word in one speech), identification of speaker (the test case made a distinction based on gender and political party, for the purpose of determining a political leaning it would be important to identify whether the person is in government or represents an oppositional party). While there are obvious problems with using this method as a means to determine political leanings; from the outset what language and issues can be considered of concern to the left or right, should be addressed. With some further consideration DiLiPaD might offer a more accurate record of a government's political position, both in general terms and on particular issues.

3. Preparing the Data

3.1. Data Acquisition

The HUB Global Labour Conflicts project (herewith, HUB) was initiated by IISH in 2015 as a collaborative effort to connect ‘information about labour conflicts through time and space’ (“HUB”). Overseen by Sjaak van der Velden, the data published through a ‘Dataverse’ repository (an open-source web application for the preservation of research data) contained over 50 different datasets. HUB is now updated yearly and managed by Rosa Kösters, a PhD researcher at the IISH. It networks together over 120 scholars to facilitate information and data exchange for expanding the corpus of the Dataverse repository in ways which encourage interdisciplinary research on labour relations (“HUB”).

The International Institute of Social History's 'HUB Global Labour Conflicts' database with URL <https://datasets.iisg.amsterdam/dataverse/labourconflicts> in a 'Dataverse' repository.

The Dutch Labour Conflicts (DLC) data is a subset of the data collected in the HUB Global Labour Conflicts Dataverse. This subset contains information on over 15,000 labour conflicts which occurred within the Netherlands between 1372 and 2019. It is presented by the IISH in the data hub 'DRUID' under a CC0 1.0 Universal licence, meaning the data is committed to the public domain and may be accessed, copied, modified, and redistributed by anyone without additional permissions, even for commercial purposes (Creative Commons).

The International Institute of Social History's Dutch 'Labour Conflicts' (1372-2019) Linked Data with URL <https://druid.datalegend.net/IISG/Labourconflicts> in data hub 'DRUID'.

DRUID is an online environment where users can 'upload, transpose, edit, query, and present' datasets with provided tools which facilitate data management and manipulation for researchers who may or may not have technical/computer science experience ("Tools"). The purpose of DRUID as a data hub (rather than a data repository) is to craft Linked Data from multiple datasets. Links are established between datasets when data records share entries within Resource Description Framework (RDF) semantic 'triples': statements which distinguish a 'subject' – 'predicate' – 'object' relationship between graphical data records (RDF Working Group). The DRUID environment includes built-in tools to index datasets using RDF syntax and to then query data records via a 'SPARQL Protocol and RDF Query Language' (SPARQL) endpoint. The RDF schema of the DLC dataset is visualised on the

following page in figure 5. An example query made on the data records using the built-in DRUID SPARQL endpoint is provided on the following page in figure 6. A single Linked Data entry from the DLC dataset in DRUID is exemplified in the following pages as figure 7.

SPARQL queries allow users to download a desired filtered data subset in the format of a .csv (Comma Separated Values) file. The .csv format is a text file which uses commas as delimiters to separate individual text records and this format, according to Shafranovich, ‘has been used for exchanging and converting data between various spreadsheet programs for quite some time’ (1). This non-proprietary file format is the most commonly used format by researchers in the humanities and social sciences (e.g. historians) as it is easily adapted for use in Microsoft Excel with a ‘minimal amount of fuzz’ (Zijdeman et al. 2).

By running a SPARQL query on the DLC dataset via DRUID, a researcher at IISH had previously collected information on labour conflicts which occurred in the Netherlands between 1950 and 2019. This data sample, presented as a .csv file and showcased in the following pages as figure 8, was privately provided for further use, cleaning, and enrichment but could be assembled by any user at the SPARQL endpoint on DRUID.

3.2. Data Cleaning

The .csv file received from IISH was first opened using a basic text editor (in this case, the free, open-source software ‘Notepad++’). Upon inspection, several parsing issues were identified in the file which would inhibit interoperability of the file in other software applications which use the comma as a delimiter between data records and employ sets of double quotes (") to enclose fields which may contain (as a part of the record itself) line breaks, double quotes, or commas (Shafranovich, 2). The data contains a column labelled ‘description’ under which the details of individual labour conflicts are stored. The text in ‘description’ data records sometimes includes commas, double quotes, and line breaks which, when not enclosed by double quotes, may cause certain software to parse the individual data

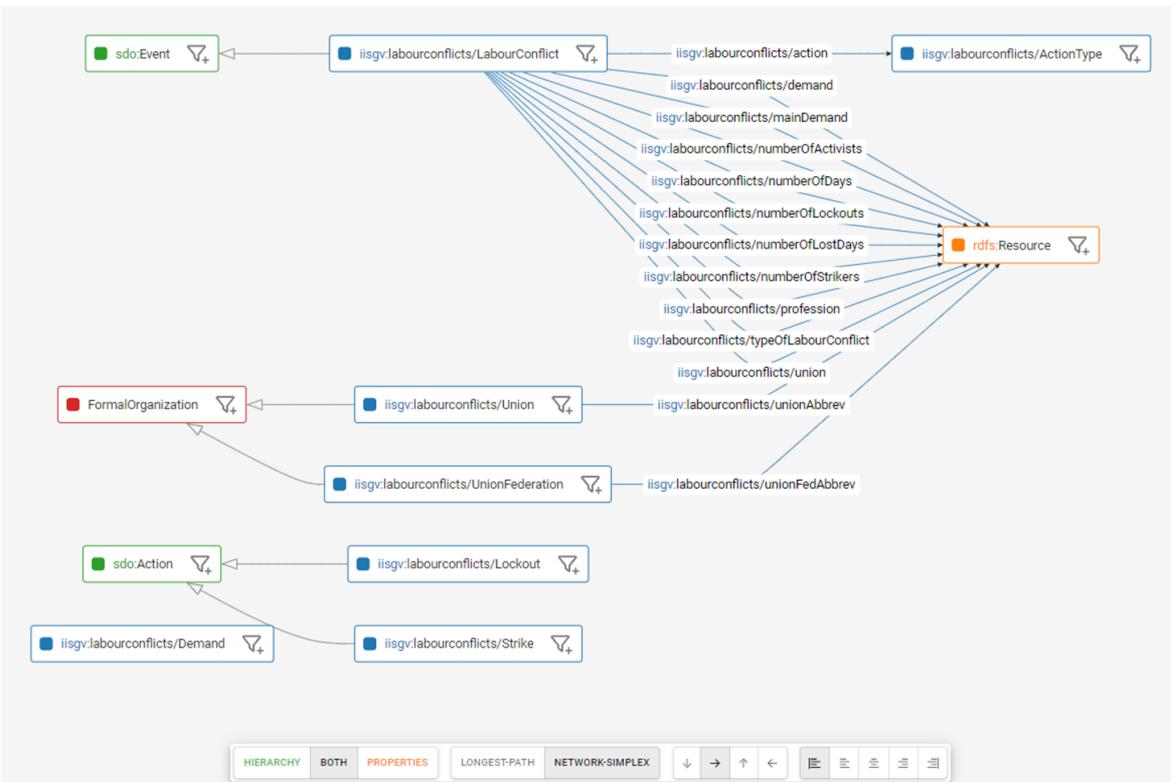


Fig. 5. IISH DLC Linked Data RDF Schema in DRUID.

```

1 PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
2 PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
3 PREFIX schema: <https://schema.org/>
4 PREFIX geo: <http://www.opengis.net/ont/geosparql#>
5 PREFIX skos: <http://www.w3.org/2004/02/skos/core#>
6 PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
7 PREFIX vocab: <https://iisg.amsterdam/vocab/labourconflicts/>
8 PREFIX time: <http://www.w3.org/2006/time#>
9 PREFIX owl: <http://www.w3.org/2002/07/owl#>
10 PREFIX org: <http://www.w3.org/ns/org#>
11 PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
12
13 select distinct ?id ?name ?place ?year
14 {
15   ?id rdf:type vocab:LabourConflict;
16   schema:name ?name;
17   schema:startDate ?startDate ;
18   schema:location/schemaname ?place.
19   BIND(REPLACE(STR(?startDate), "(\d{4})(-\d{2}-\d{2})", "$1") AS ?year)
20   FILTER regex(?year, ?theyear, 'i')
21   FILTER regex(?place, ?thelocation, 'i')
22   FILTER(LANG(?name) = "" || LANGMATCHES(LANG(?name), "nl"))
23 }

```

Fig. 6. Example SPARQL endpoint query of the IISH DLC Linked Data which selects all labour conflicts occurring in the year '1956' in the place 'Rotterdam'.

<https://iisg.amsterdam/id/labourconflicts/12782>

Strike, - 1950

Labour Conflict

<https://iisg.amsterdam/id/labourconflicts/12782>

Type	Labour Conflict
Type of action	Strike
Strike days	125
Number of strikers	25
Occupations involved	Leatherworks (unspecified)
Agent	Onbekend Unknown
Name	Staking, - 1950 Strike, - 1950
StartDate	1950-01-01

Fig. 7. Example (id# 12782) of IISH DLC Linked Data in DRUID.

```

1 "id","","startdate","","duration","","description","","character_nl","","character_en","","type_of_strike_nl","","t^
2 "00003,""2001-12-20""",""0.1"",""Tientallen medewerkers legden enige tijd het werk neer, omdat er als gevol
3 "00004,""2001-04-11""",""2"",""De eerste dag legden 900 mensen het werk neer en een week later nog eens 120
4 "00005,""2001-05-18""",""6"",""Het personeel was gefrustreerd door de reeks van organisaties en maakte zich
5 "00006,""2001-12-14""",""0.5"",""Een collega van lijn 33 had van een passagier klappen gehad. Na het protes
6 "00007,""2001-12-04""",""1"",""Om te bezuinigen wilde de directie dat de bemanningen lading zelf zouden gaa
7 "00008,""2001-04-23""",""0.5"",""Voor een nieuwe CAO voor de 50.000 betrokken werknemers van de technische
8 "00009,""2001-12-03""",""0.5"",""Honderden deden aan de actie tegen de door de Finse directie opgelegde slu
9 "00010,""2001-03-12""",""3"",""In het onderwijs werd drie maal een halve dag gestaakt voor meer geld voor h
10 "00011,""2001-03-02""",""31"",""De bonden eisten een loonsverhoging van twee maal vier procent en twee keer
11 "00012,""2001-05-22""",""2"",""De medewerkers van het distributiecentrum van deze speelgoedgroothandel met
12 "00013,""2001-04-05""",""3"",""Een vervolg op de acties in 2000 tegen Bestemming Klant. Op 9 januari kondig
13 "00014,""2001-01-18""",""0.2"",""De mensen hadden in 2000 geen loonsverhoging gekregen en stelden op 16 jan
14 "00015,""2001-05-29""",""0.4"",""Een collega was door een passagier mishandeld en daarom legde het persone
15 "00016,""2001-09-25"",""1"",""Een protest tegen de plannen van Europa om reders de vrije hand te geven bij
16 "00017,""2001-09-14"",""1"",""Het personeel van de operatiekamer en de Verkoeverkamer onderbrak het werk om
17 "00018,""2001-10-30""",""0.5"",""Voor een loonsverhoging van 6-7%. Bovendien 'bestormden' de stakers het st
18 "00023,""2001-05-07""",""32"",""Voor een loonsverhoging van 4% en tegen verlenging van de werkweek van 36 n
19 "00024,""2001-11-15""",""3"",""De bonden eisten een loonsverhoging van 7%, incl. een dertiende maand voor d
20 "07489,""1981-04-13""",""14"",""Tegen onderbetaling. De havenarbeiders van Seaport Terminals weigerden uit
21 "10476,""1961-11-29"",""1"",""De meisjes eisten loonopslag of ontslag."",""Wild"",""Wildcat"",""Klassiek"
22 "10633,""1996-03"",""6"",""De bezetters deelden op straat krentenbollen uit, maar het mocht niet baten."",
23 "10634,""1972-10-18"",""0.5"",""Onbekend"",""Unknown"",""Klassiek"",""Classical"",""Onbekend/onbeslist
24 "10635,""1972-10-27"",""0.1"",""Wild"",""Wildcat"",""Klassiek"",""Classical"",""Onbekend/onbeslist"|,
25 "10636,""1972-11-09"",""0.5"",""Wild"",""Wildcat"",""Klassiek"",""Classical"",""Verlies"",""Lost"",""
26 "10637,""1972-11-28"",""0.5"",""Op 4 december ook te Zwolle, Amsterdam, Groningen en Nijmegen na de eerste
27 "10638,""1972-11-29"",""1"",""Wild"",""Wildcat"",""Klassiek"",""Classical"",""Verlies"",""Lost"",""St
28 "10639,""1972-11-30"",""0.3"",""Wild"",""Wildcat"",""Werkonderbreking"",""Walkout"",""Onbekend/onbeslis

```

Fig. 8. Head (27) of .csv file provided by IISH from the DLC containing information on labour conflicts in the Netherlands between 1950 and 2019.

records in a manner which does not accurately represent the data record. In addition, it was noted that the .csv file did not consistently represent line breaks and employed (uneven) strings of semicolons to signify new lines. To address issues in the .csv file, the following steps were taken within the text editor using the ‘Find’ and ‘Replace’ tools:

1. Find and replace character strings ';;;' and ';;' with '' (remove).
2. Find and replace character strings '";' and ';"' with ';'.
3. Add '''' character to the beginning of Lines:
2068, 2070, 2075, 2078, 2080, 2082, 2084-2085, 2087-2089, 2091-2092,
2095, 2104, and 2109-2110.
4. Edit to remove 'newlines' (* as lines are removed, line numbers do change) in the beginning of Lines:
266, 315-324, 551-553, 747, 766-781, 798-800, 817, 1007, 1172-1174,
2905-2906, 2992-3009, 2994-2996, 2996, 2997, 2998-2999, 3002, 3009,
3018-3019, 3019-3020, 3024, 3032, 3049, 3060, 3069-3070, 3072, 3082,
3115, 3131, 3198, and 3257.
5. Find and replace character strings '"""', '''' and ',''' with ','.
6. Find and replace character strings '""'''', '""''', and '""'' with ''''.

The .csv file then displayed 3,334 rows of data entries related to individual labour conflicts in the Netherlands between 1950 and 2019. However, column parsing errors did remain between data records in many rows. In a text editor, these errors are difficult to identify and manipulate in large scale and so, the .csv file was saved as a copy and opened using the open-source desktop application ‘OpenRefine’ (version 3.7.7), a ‘tool for working with messy data’ (OpenRefine, “OpenRefine”).

When creating the new project in OpenRefine from a copy of the .csv file, it was necessary to specify that the file used character encoding ‘UTF-8’ (Unicode Transformation Format – 8-bit; an ISO standard encoding language for electronic communication), that data columns were to be separated using commas (rather than tabs or a custom character), that the

first line was to be parsed as column headers, that the application should not trim the leading or trailing whitespaces in data records (whitespaces are typographical characters which represent vertical or horizontal spaces in text), and that the application should not use a character (e.g. a double quote) to enclose cells containing the column separator (a comma).

OpenRefine offers data transformation using GREL (the General Refine Expression Language), a language designed to resemble Javascript by using formulas and variables as expressions (OpenRefine, “General Refine Expression Language”). It also offers tools for text filtering and the faceting of text, numbers, timelines, or custom facets. Using these tools, the first three columns of data records – ‘id’, ‘startdate’, ‘duration’ – were found to display the text of data records accurately and consistently. Parsing errors began to appear in the fourth column (‘description’) and to manage these issues, the following actions were taken:

1. Text facet column 'description', select (blank), invert selection.
2. Transform faceted column 'description' with GREL using expression {
 '@' + value }*.
 * the '@' character was chosen because none exist in the .csv
3. Remove text filter.
4. Using text facets and filters when appropriate, locate the end of the record for each (non-blank) record belonging to column 'description'. Transform faceted/filtered columns with GREL using expression { value + '@' }. Invert or remove text facets and filters when necessary.
5. After performing a transformation on column 'province_7', 25 single cell edits were manually performed for efficiency.

An extract of these manipulations as .json file ‘COPY1_operation_history.json’ with URL
https://github.com/Preservation-and-Access-2023/Labour-Conflicts/blob/main/History/COPY1_operation_history.json in a ‘Github’ repository folder.

The OpenRefine project was then exported as a second copy of the .csv file and opened again with the text editor to perform the following steps:

1. Find and replace character strings '@'', and ''@' with '@'.
2. Find and replace character string '""" with '''.

The second copy was saved in the text editor and opened as a new project in OpenRefine, specifying for the application to distinguish pairs of '@' characters as enclosing a record containing the comma delimiter. Using the text faceting tool, errors were located in the ‘description’ column within Lines 880, 2792, 2842, 3085, and 3266. Errors were addressed in the text editor and the second .csv copy was saved and reloaded again as a new project in OpenRefine. The technique of viewing the data in OpenRefine whilst addressing errors in the text editor was repeated until all necessary data records were enclosed with pairs of '@' characters. For example, the ‘place’ of ‘Warga, Leeuwarden’ or the ‘occupation_nl’ of ‘vleeswaren, slachterij’ / ‘occupation_en’ of ‘Food and beverages (meat products, abattoir)’ – all of which contain the comma delimiter within the data record. When the dataset was fully cleaned and inspected in OpenRefine, it was exported as a new .csv file. Based on the .csv file standard, the enclosing ‘@’ pairs revert to double quotes for improved interoperability.

Cleaned data subset ‘CLEAN_Stakingen-Nederland-1950-2019.csv’ with URL
https://github.com/Preservation-and-Access-2023/Labour-Conflicts/blob/main/History/CLEAN_Stakingen-Nederland-1950-2019.csv in a ‘Github’ repository folder.

3.3. Data Enrichment

To enhance the cleaned data for the purpose of adding political context, it was necessary to collect information regarding the political structures and climate of the Netherlands within the temporal scope of the dataset (1950-2019). Contributors to Wikipedia, The Free Online Encyclopedia, had graphically assembled the desired data records on the page ‘List of

cabinets of the Netherlands' (seen below in figure 9). The complete list contains information on Dutch political cabinets since 1877 and is stored as a Wikimedia list article in Wikidata (Q3178161). A Python code was written to scrape, clean, format, and filter the data before exporting it to a .csv file for later reference.

List of cabinets of the Netherlands

From Wikipedia, the free encyclopedia

Main article: [Cabinet of the Netherlands](#)

The following is a list of cabinets of the Netherlands since 1877.

List [edit]

Cabinet	Prime Minister	Term of office		Demissionary	Time in office	Parties	Political position	Orientation	Legislature Status	Type	Election
Kappeyne van de Coppelio [1]	Jan Kappeyne van de Coppelio	3 November 1877	20 August 1879	11 July 1879	1 year, 290 days	[IL]	Centre-right	Liberal	Majority	Conventional	1877
Van Lynden van Sandenburg [2]	Theo van Lynden van Sandenburg	20 August 1879	23 April 1883	1 March 1883	3 years, 246 days	[ICO] • [IL] • [ICA] [IP]	Right-wing	Christian Democratic-Liberal	Minority	Confidence and supply	1879
J. Heemskerk [3]	Jan Heemskerk	23 April 1883	21 April 1888	30 March 1888	4 years, 364 days	[ICO] • [IL] • [ICA]	Right-wing	Liberal	Minority	Confidence and supply	1883
Mackay [3]	Aeneas Mackay	21 April 1888	21 August 1891	9 July 1891	3 years, 122 days	ARP • [ICA] • [IP]	Centre-right	Christian Democratic	Minority	Confidence and supply	1888
Van Tienhoven [3]	Gijsbert van Tienhoven	21 August 1891	9 May 1894	24 April 1894	2 years, 261 days	LU • [IL]	Right-wing	Liberal	Majority	Conventional	1891
...											

Fig. 9. Head (5) of the ‘List of cabinets of the Netherlands’ from Wikipedia.

Python code for collecting data on Dutch political cabinets ‘cabinets_collector.py’ with URL https://github.com/Preservation-and-Access-2023/Labour-Conflicts/blob/main/History/cabinets_collector.py in a ‘Github’ repository folder.

Dutch political coalition cabinets (1950-2019) dataset ‘Dutch_cabinets.csv’ with URL https://github.com/Preservation-and-Access-2023/Labour-Conflicts/blob/main/History/Dutch_cabinets.csv in a ‘Github’ repository folder.

The columns of the cleaned DLC dataset are described and graphically showcased on the following page as table 1. The dataset was expanded using OpenRefine where a blank column labelled ‘political_cabinet’ was added to the end of the rows. Viewing the ‘List of cabinets’ for start and end dates of political terms of service whilst utilising text facetting on the ‘startdate’ column in OpenRefine, data entries (rows in the ‘Dutch_cabinets.csv’ file) were copied into the blank column for all 3,334 rows based on dates. When the column was complete, the following actions were taken to adapt the new data to the existing file:

1. Split column ‘political_cabinet’ into several columns by separator ‘,’ and select to remove this column.
2. Rename column ‘political_cabinet 1’ to ‘political_cabinet’.
3. Remove columns ‘political_cabinet 2’ and ‘political_cabinet 3’.
4. Move column ‘political_cabinet 4’ to the end.
5. Rename column ‘political_cabinet 5’ to ‘political_position’.
6. Split column ‘political_cabinet 4’ into several columns by separator ‘•’ and select not to remove this column.
7. Replace strings ‘•’ with ‘, ’ in column ‘political_cabinet 4’.
8. Rename column ‘political_cabinet 4’ to ‘political_coalition’.
9. Rename columns ‘political_cabinet 4 #’ to
‘politicalparty_abbreviation_#’ (where # = 1 to 5).
10. Add 2 columns based on column ‘politicalparty_abbreviation_#’, named
‘politicalparty_nl_#’ and ‘politicalparty_en_#’ (where # = 1 to 5).
11. Using text facetting, expand political party abbreviations in columns
‘politicalparty_nl_#’ and ‘politicalparty_en_#’ (where # = 1 to 5) to formal political party names (accordingly in [Nederlands/English]).

Formal names for the political parties of the Netherlands in both Dutch and English were sourced by navigating Wikipedia/Wikidata pages via hyperlinks from the ‘List of cabinets’ page. This choice enhances interoperability of the .csv file as these values can be reconciled

Table 1. Column Label Names and Descriptions of the DLC Data Subset

Label Name	# of Repeats	Label Description
1 id		Data entry identification number.
2 startdate		Start date of labour conflict; expressed as YYYY-MM-DD (if the exact date is unknown, the expression is limited to YYYY-MM or YYYY).
3 duration		Duration of labour conflict; expressed in calendar days (inclusive of Sundays and bank holidays); if less than 1 day, expressed using 0.1 day = 1 hour.
4 description		Description or details of labour conflict event.
5 character_nl		Character of labour conflict; with ('Vakbond' / 'Union') or without ('Wild' / 'Wildcat') union leadership. [Nederlands / English]
6 character_en		
7 type_of_strike_nl		Type of strike action during labour conflict. [Nederlands / English]
8 type_of_strike_en		
9 result_nl		Result of labour conflict. [Nederlands / English]
10 result_en		
11 action_nl		Strike action; ('Staking' / 'Strike' ; 'Uitsluiting' / 'Lockout' ; 'Geen actie' / 'Nothing happened'). [Nederlands / English]
12 action_en		
13 place_#	132	Location (city; village; region; etc.) where labour conflict occurred.
14 province_#		Dutch province where the labour conflict occurred.
15 company_name_#	78	Name of company where the labour conflict occurred.
16 company_place_#		Location (city; village; region; etc.) where the company is conducting business.
17 company_province_#		Dutch province where the company is conducting business.
18 occupation_hisco_#	8	Historical International Standard of Classification of Occupations (HISCO) code for occupation (see: https://historyofwork.iisg.nl/).
19 occupation_nl_#		Name (and/or character) of occupation. [Nederlands / English]
20 occupation_en_#		
21 TotalStakers		If the entry describes a strike, total number of striking labourers.
22 TotalGestaakteDagen		If the entry describes a strike, total number in person-days of labour lost.
23 TotalUitgeslotenen		If the entry describes a lockout, total number of locked out labourers.
24 TotalVerlorenDagenUitgeslotenen		If the entry describes a lockout, total number in person-days of labour lost.
25 TotalAantalActievoerders		If the entry does not describe a strike or lockout, total number of participants.
26 andere_actie_nl_#	4	Additional actions taken during the labour conflict. [Nederlands / English]
27 andere_actie_en_#		
28 labourunion_abbreviation #	4	Abbreviation of the name of the involved labour union.
29 labourunion_name_#		Name of the involved labour union.
30 labourunionfederation_abbreviation #		Abbreviation of the name of the involved labour union federation.
31 labourunionfederation_name_#		Name of the involved labour union federation.
32 reason_#	4	Alphanumeric code related to the reasons for / demands of the labour conflict.
33 main_reason_#		Whether the reason was the main reason for the labour conflict; ('no' ; 'yes').
34 demand_nl_#		
35 demand_en_#		Explanation of the demand from the striking labourers. [Nederlands / English]

against Wikidata services in OpenRefine. The following steps were taken to include URLs of the Wikidata items based on the names of the Dutch political cabinets:

1. Add column named 'recon', based on column 'political_cabinet' with GREL using expression { value.split(' ')[0] + ' cabinet' }.
2. Reconcile column 'recon' with service 'Wikidata reconcili.link (en)' to type 'Cabinet of the Netherlands (Q2479200)'.
3. Match all data records in column 'recon' with Wikidata entries.
4. Add entity identifiers column named 'political_cabinet_wikidata' based on column 'recon'.
5. Transform column 'political_cabinet_wikidata' with GREL using expression { 'https://www.wikidata.org/wiki/' + value }.
6. Remove column 'political_cabinet'.
7. Rename column 'recon' to 'political_cabinet'.

The OpenRefine project ‘Dutch-Labour-Conflicts-and-Politics_1950-2019’ then included seven new columns (exemplified below as table 2) and was exported as a .csv file. It was uploaded as user ‘Preservation-and-Access 2023’ to a ‘Github’ repository called ‘Labour-Conflicts’ alongside a folder titled ‘History’ containing the other relevant files listed within this paper. This presentation of enriched data from the DLC subset is titled ‘Dutch Labour Conflicts and Politics (1950-2019)’ (herewith, DLCP) and is available publicly.

Table 2. Column Label Names and Descriptions as Added to the DLC Data Subset

	Label Name	# of Repeats	Label Description
36	political_cabinet		Name of political cabinet; derived from the surname of the serving Prime Minister.
37	political_cabinet_wikidata		URL to corresponding Wikidata entry for each ‘political_cabinet’.
38	political_position		Political leaning of the serving political cabinet.
39	political_coalition		Abbreviations of political parties participating in the serving political coalition.
40	politicalparty_abbreviation_#	5	Abbreviation of an individual political party serving in the political coalition.
41	politicalparty_nl_#		Full name of an individual political party serving in the political coalition. [Nederlands / English]
42	politicalparty_en_#		

DLCP data subset ‘Dutch-Labour-Conflicts-and-Politics_1950-2019.csv’ with URL https://github.com/Preservation-and-Access-2023/Labour-Conflicts/blob/main/Dutch-Labour-Conflicts-and-Politics_1950-2019.csv in a ‘Github’ repository.

The DLCP enriched data subset is by no means a complete practice for enhancing the data available from the IISH. The steps taken here are replicable and repeatable for data subsets which might be extracted from other SPARQL queries via the DRUID endpoint. The enhancements applied from reconciling data records with Wikidata entries could also be practised on other columns within the DLCP to expand the dataset further. The work performed during this project is merely one example of the potential benefits and limitations of Linked Data usage and exploration.

4. Reflections on Data Usage

One major hurdle for understanding the data in the DLC subset was a lack of documentation explaining and describing the labels (column headers) used to separate data records. No metadata with column descriptions were provided with the DLC database, so the descriptions offered in table 1 were based on those in Appendix 1 of Van der Velden’s dissertation (313-18). Without provided definitions (in an easy-to-find location) for data records or data labels, the user of the data is prone to interpret or assume what these fields mean. A .json metadata file containing all label names and descriptions (tables 1 and 2 from above) of the DLCP included in the Labour Conflicts Github repository can draft insight for users which align better with the intent of the data provider (CSV on the Web Working Group). By including a URL hyperlink for Wikidata entries on Dutch political cabinets in the DLCP enriched data, the intention was to provide users an opportunity to explore data records further and gather more information from sources shared by other users, providers, and

researchers via Linked Data. The enrichment of the DLC with data from the ‘List of cabinets’ also gives users and data providers new options for visualising the data in various ways – for, as Kieran Healy wrote, ‘good visualization methods offer extremely valuable tools that we should use in the process of exploring, understanding, and explaining data’ (ch.1). The DRUID environment and IISH data story are already prepared with visualisation tools for managing these requests.

Data labels and descriptions as Microsoft Excel file ‘Labels_and_Descriptions.xlsx’ with URL https://github.com/Preservation-and-Access-2023/Labour-Conflicts/blob/main/Labels_and_Descriptions.xlsx in a ‘Github’ repository.

In order to visualise the DLCP, figure 10 (on next page) shows a histogram of the average strike frequency for each cabinet which held office between 1950 and 2019. The bars corresponding to these cabinets are coloured according to their political positions as given in the ‘List of cabinets of the Netherlands’ Wikipedia page. This visualisation illustrates some pitfalls of using this dataset. Since there has been no centre-left or left-wing cabinet in the Netherlands after the First Kok cabinet took office in 1994, any analysis of the impact of left- or right leaning cabinets should be constrained to the period 1950-1994, since there is no data on how a left-leaning cabinet would function in the Dutch socio-economic climate of the 21st century. Of course, such a visualisation shows only a limited number of facets of the data.

After adding the political data, the dataset could be converted to Linked Open Data. It would then be possible to link the DLCP to Wikidata and other Linked Open Data, which would in theory largely expand the number of possible facets of analysis of the data and give more researchers the opportunity to work with this dataset. However, the usefulness of linking to cabinets as Wikidata objects is greatly reduced by the inconsistencies in Wikidata. For example, Wikidata objects of recent cabinets like the Third Rutte cabinet (Q42293409)

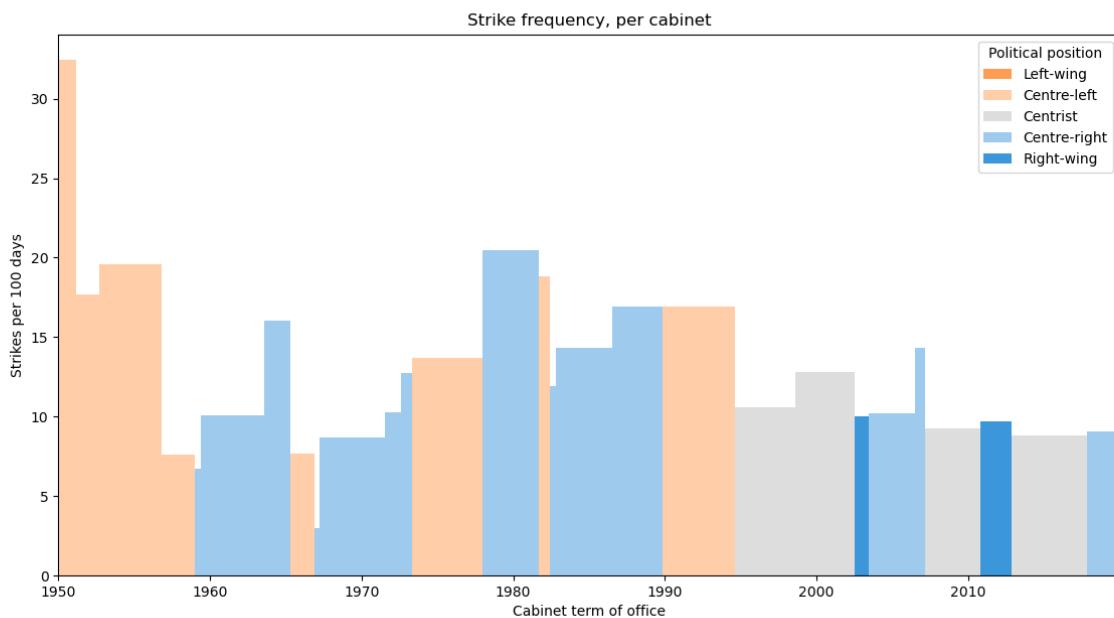


Fig. 10. A visual representation of the DLCP dataset: the average frequency of strikes (expressed in number of strikes per 100 days) for each Dutch cabinet in the period 1950-2019. Bars are coloured according to the political position of the cabinet.

use the ‘has parts’ property (P527) to list all ministers and secretaries of state that were part of the cabinet, while earlier cabinets such as the Drees/Van Schaik cabinet (Q1719852) only use the ‘head of government’ property (P6) to link to the prime minister, in this case Willem Drees (Q12769). These inconsistencies make it difficult to do comparative analysis. Other data that is recorded in the ‘List of cabinets’ on Wikipedia is not represented on Wikidata at all: this includes the coalition parties, political position, political orientation, legislature status, type, and election. If such data were represented as links in Wikidata, simply linking strikes to the cabinet Wikidata object would be much more powerful as it would open the door to analysis of all of these factors. As things are now, it is necessary to explicitly incorporate data such as the political position in the DLCP database. Another possibility would have been to add these links to Wikidata, for example using the ‘political alignment’ property (P1387) to link cabinets to the political positions ‘centre-left’ (Q737014), ‘centrism’

(Q6587194), ‘centre-right politics’ (Q844072) and ‘right-wing’ (Q76074). However, a SPARQL query of Wikidata shows that as of yet no cabinet has been given the ‘political alignment’ property, presumably because such labels are likely to be contentious.

5. Conclusions

The DLC dataset was started by Sjaak van der Velden in 2000 and later developed at IISH using data aggregated by the CBS in combination with strikes mentioned in newspapers, academic literature, and trade union periodicals. Despite the general term ‘labour conflicts’ the DLC dataset refers mostly to unionised strikes, effectively excluding other types of workers’ resistance – walkouts, sabotages and boycotts – and different forms of labour – unpaid domestic work and the labour of enslaved people.

A potential question for a political historian using the dataset could be; is there any correlation between the political leaning of the government and the number and intensity of strikes? Within the context of this question, some consideration was given to how the political leaning of a government might be classified, with DiLiPaD being mentioned as a potentially more accurate method to determine the political leaning of a government.

The DLC dataset has the potential to link to other Linked Open Data. However, the messiness of the .csv file and the lack of metadata describing the data provenance and column descriptions make it difficult to interpret the data. Having cleaned the .csv and reconstructed the metadata based on Van der Velden’s dissertation, it was possible to link the data to data about Dutch cabinets. The thusly formed DLCP has the potential to help political scientists and political historians better understand the relations between strikes and national politics in the Netherlands. The usefulness of further linking this data to Wikidata is limited due to the inconsistency and incompleteness of Wikidata.

One area in which further work is needed is in transforming the DLCP into Linked Open Data. At the same time, Wikidata objects representing Dutch cabinets could be linked

to other Wikidata objects, such as the political orientation and coalition parties, to allow for more complex and interesting SPARQL queries. In order to better place the relation between strikes and politics, it could also be useful to consider data from other geographical areas and time periods. While this research has focused specifically on the DLC, and even then only on a specific time period, similar research could be conducted more broadly. For example, other historical time frames in the Netherlands could be considered and a comparison between those conclusions and the conclusions made in this paper could be made. This research could furthermore be conducted on an international scale, using other labour conflict datasets made available through HUB. This would be helpful for continuing this research by making variables such as the political orientation of relative governments easier to isolate.

Works Cited

- Alimahomed-Wilson, Jake, and Ellen Reese. "Surveilling Amazon's Warehouse Workers: Racism, Retaliation, and Worker Resistance amid the Pandemic." *Work in The Global Economy*, 1.1 (2021): 55–73. Web. *Bristol UP*.
- Aurich, Jens, Catherine Simpson, and Rosa Kösters. "Workers' Resistance across Time and Space: Presenting the HUB Global Labour Conflicts." *Arbetarhistoria* 20.2 (2022): 161-75.
- Beelen, Kaspar, Jonathan Blaney, and Luke Blaxhill. "Jisc Final Report – Digging into Linked Parliamentary Data (DiLiPaD)." (2016). repository.jisc.ac.uk/6544/.
- Blaney, Jonathan. "Introduction to the Principles of Linked Open Data." *Programming Historian*, last modified 1 December 2021, <https://doi.org/10.46430/phen0068>.
- Blaxhill, Luke, and Kaspar Beelen. "A Feminized Language of Democracy? The Representation of Women at Westminster since 1945." *Twentieth Century British History* 27.3 (2016): 412–49.
- Coates, Ta-Nehisi. "The Case for Reparations." *The Atlantic*, June 2014, www.theatlantic.com/magazine/archive/2014/06/the-case-for-reparations/361631/.
- Creative Commons. "CC0 1.0 Deed: CC0 1.0 Universal." *Creative Commons*, creativecommons.org/publicdomain/zero/1.0/. Accessed 15 December 2023.
- CSV on the Web Working Group. "CSV on the Web: A Primer." *W3C Working Group Note*, 25 February 2016, www.w3.org/TR/2016/NOTE-tabular-data-primer-20160225/.
- Federici, Silvia. *Revolution at Point Zero: Housework, Reproduction, and Feminist Struggle*. Second edition, [Revised and Expanded edition]. Oakland, CA: PM Press, 2020.
- Healy, Kieran. "Look at Data". *Data Visualization: A Practical Introduction (Incomplete draft)*, Princeton University Press, 25 April 2018, ch. 1, socviz.co/lookatdata.html.

“HUB Global Labour Conflicts.” *Blog*, International Institute of Social History, 28 February 2022, iisg.amsterdam/en/blog/hub-global-labour-conflicts.

International Institute for Social History. “Labour Conflicts Database Schema.”

Labourconflicts, IISG, DRUID, 2021,

druid.datalegend.net/IISG/Labourconflicts/schema?a=UL.

———. “Search for strikes in Labour Conflicts database [Zoeken naar stakingen in Labour Conflicts database].” *Stories*, IISG, DRUID, druid.datalegend.net/IISG/-/stories/Querying-the-Labour-Conflicts-database. Accessed 15 December 2023.

———. “Strike, – 1950.” *Labourconflicts*, DRUID, 2021, druid.datalegend.net/IISG/Labourconflicts/browser?resource=https://iisg.amsterdam/id/labourconflicts/12782.

OpenRefine. “General Refine Expression Language.” *User Manual*, Open Refine, last updated 24 November 2022, openrefine.org/docs/manual/grel.

———. “OpenRefine.” *OpenRefine*, openrefine.org. Accessed 17 December 2023.

RDF Working Group. “RDF 1.1 Concepts and Abstract Syntax.” *W3C Recommendation*, 25 February 2014, www.w3.org/TR/2014/REC-rdf11-concepts-20140225/.

Shafranovich, Yakov. “Common Format and MIME Type for Comma-Separated Values (CSV) Files.” *The Internet Society*, October 2005, datatracker.ietf.org/doc/html/rfc4180.

Star, Susan Leigh. “Power, Technology and the Phenomenology of Conventions. On Being Allergic to Onions.” *The Politics of Interventions*, edited by Kristin Asdal and Brita Brenna, Unipub, 2007, pp. 79–104.

“Tools – dataLegend.” *Tools*, DataLegend, www.datalegend.net/tools/. Accessed 15 December 2023.

Velden, Sjaak van der. *Stakingen in Nederland: Arbeidsstrijd 1830-1995*. Diss. Universiteit Leiden, 2000. Amsterdam: Stichting Beheer IISG/NIWI, 2000. Print.

- . “Strikes, Lockouts, and Informal Resistance.” *Handbook the Global History of Work*, edited by Marcel van der Linden and Karin Hofmeester, De Gruyter Oldenbourg, 2019, pp. 521-550.
- Wikipedia contributors. “Left-right political spectrum.” *Wikipedia, The Free Encyclopedia*, last modified 11 Dec. 2023, en.wikipedia.org/wiki/Left-right_political_spectrum.
- . “List of cabinets of the Netherlands.” *Wikipedia, The Free Encyclopedia*, last modified 19 Nov. 2023, en.wikipedia.org/wiki/List_of_cabinets_of_the_Netherlands.
- . “List of political parties in the Netherlands.” *Wikipedia, The Free Encyclopedia*, last modified 14 Dec. 2023, en.wikipedia.org/wiki/List_of_political_parties_in_the_Netherlands.
- Zijdeman, Richard, et al. “IAP 2019 report on CLARIAH Structured Data (WP4): dataLegend.net.” *dataLegend*, 29 March 2019, docs.google.com/document/d /1U_gHHm6np3IlKOih-VWYpuByHFCUrpsNcnMxT127G8.