Journal of Open Humanities Data

Mapping Historical Detective Fiction Worldwide: A Bibliographic Dataset of 301 Novels and Series

DATA PAPER

MARINA BULAKHOVA

ubiquity press

ABSTRACT

This annotated bibliographic dataset describes 301 historical detective novels and series first published in 46 countries and 36 languages; 170 entries are non-Anglophone. The criteria for inclusion in this dataset was as follows: at least fifty years between publication and earliest time of action; recognized detective status or distinctive detective plot structure; and an adaptation or two of four further markers of significance (awards, scholarly attention, translations, reissues). Candidates were gathered with large language models, then selected and verified by the author against cited sources. The data are deposited on Zenodo as CSV and XLSX files and can be reused in crime fiction studies and comparative genre history.

CORRESPONDING AUTHOR:

Marina Bulakhova

Independent Researcher, Tartu, Estonia

marinabulakhova@gmail.com

KEYWORDS:

historical fiction; detective fiction; historical mystery; genre studies; comparative literature; bibliographic dataset

TO CITE THIS ARTICLE:

Bulakhova, M. (2026). Mapping Historical Detective Fiction Worldwide: A Bibliographic Dataset of 301 Novels and Series. *Journal of Open Humanities Data*, 12: 149, pp. 1–6. DOI: https://doi.org/10.5334/johd.676

(1) OVERVIEW

REPOSITORY LOCATION

Version v1 https://doi.org/10.5281/zenodo.22858682, Version v2 https://doi.org/10.5281/zenodo.23373266.

CONTEXT

Historical detective fiction is a well-established branch of popular fiction with dedicated scholarship (Browne & Kreiser, 2000; de Groot, 2009). Its bibliographic resources, however, are either restricted to English-language publication, like Hubin's (2003) bibliography of crime fiction, or compiled for readers rather than researchers, like the historical period index of Stop, You're Killing Me! (http://stopyourekillingme.com/HistoryCats/index.html), which lists titles without a full account of their setting and is likewise largely Anglocentric.

This dataset (Bulakhova, 2026) is meant as a starting point for asking how historical detective fiction maps the human past. It brings works from 46 countries into a single table in which each entry states, with sources, why the work is historical fiction, why it is detective fiction, and why it is significant enough to be included. To the author's knowledge, no other existing bibliography of the genre covers as many countries and languages. Like other curated literary datasets defined by explicit selection criteria (e.g., Grossman et al., 2022; Piper et al., 2025), it is designed as a reference sample for studying literature at the macro level, in the tradition of Moretti (2005).

(2) METHOD

STEPS

The dataset was built in June–September 2026 in three stages: candidate search, entry compilation and verification (see Quality control).

Candidate search

Candidates were proposed by large language models working under the author's instructions (see AI Declaration) and were admitted or rejected by the author against the criteria set out under Sampling strategy. The search followed four routes, summarised in Table 1.

The first round (June 2026) produced a little over two hundred entries; a second round in August extended the set to 301.

The prompts typically stated and repeated the inclusion criteria as they stood at the time, named the country (-ies), language or period to be searched, attached the current table so that only new works would be proposed, and instructed the model not to invent titles and to report when no work qualified; the models' doubts on genre attribution were treated as a signal that the author must intervene and make a decision.

Table 1 Routes of the candidate search.

ROUTE	PROCEDURE AND SOURCES
Region and country	Searches were run by cluster: the Anglophone countries; the post-Soviet states; the rest of Europe; the Middle East and Turkey; India and South-East Asia; East Asia; Africa; Latin America. A separate sweep covered early settings (before AD 1000, or outside the Greco-Roman world). A second pass went country by country, with searches run in each country's languages.
Awards	Longlists and shortlists of historical crime fiction awards (e.g., the CWA Historical Dagger, the Sue Feder Memorial Historical Mystery Award, the Agatha Award for Best Historical Novel, the Bruce Alexander Memorial Historical Mystery Award, the Herodotus Award) and of national genre prizes were worked through until new names made up less than ten percent of a list.
Indexes and adaptations	The historical period index of Stop, You're Killing Me! and Hubin (2003), including its film and television indexes, were searched. Adaptations were then traced back to their source books through Wikidata SPARQL queries, IMDb, national film databases and BBC Genome.
Language-specific sources	Sources such as ja.wikipedia and torimono.jp for Japanese *torimonochō*, Histo-Couch and Krimi-Couch for German, Hislibris for Spanish, Douban for Chinese, and the relevant categories of national Wikipedias.

DOI: 10.5334/johd.676

Entry compilation

Each entry occupies one row of thirteen columns (see Dataset Description). The commentary on genre attribution was written by the author, in some cases on the basis of initial drafts produced by language models; the column on scholarly attention was compiled by the author manually. Entries follow fixed conventions: author names are given in given-name–surname order throughout; titles appear in their published English translation or, where none exists, in the original script with a literal translation in square brackets; and a criterion that was not searched is marked with the uniform statement “not used as a criterion”.

The table was kept as a spreadsheet (Excel); batch edits and consistency checks were made with Python scripts (openpyxl) and cited web pages were archived and checked with PowerShell scripts using the Internet Archive’s Wayback Machine APIs.

SAMPLING STRATEGY

A work was included if it met three conditions.

Historical

At least fifty years had to separate the work’s first publication from the earliest time of its action. Institutions in the field set thresholds from thirty-five years (Historical Writers’ Association, 2025) through fifty (Crime Writers’ Association, 2026; Johnson, 2002) to sixty (Walter Scott Prize for Historical Fiction, 2025); fifty was adopted as a compromise. Where a work had several time layers, the earliest counted. A series was included as a whole if its first book qualified.

Detective

A settled genre label in the work’s national tradition (e.g., *giallo*, *polar*, *Krimi*, *gong’an*) or attribution to the genre by an award jury was accepted as sufficient. For hybrid works with neither, a structural test was applied: the text must contain a crime, an investigation unfolding before the reader, and an investigator. These three positions are shared by Todorov’s (1977) two-story model and by the opening of Cawelti’s (1976) formula for the classical detective story. Whether these elements are present, and how prominent they are in the plot, was judged by the author. Works without investigation *per se* (narrated as the memoirs of criminals, for example) were excluded unless a national tradition treats them as historical mystery. Spy fiction was excluded where the offense that opens the plot is directed against state security rather than being an ordinary crime.

Significant

An adaptation of any kind, including unrealized adaptation rights sold to a named buyer, was sufficient on its own. Otherwise a work needed any two of the following: a genre award, a mention in scholarly literature, translation into three or more languages, and a reissue at least twenty years after first publication. In total, 195 entries qualified through adaptation, and 106 through two or more of the other criteria.

QUALITY CONTROL

Four full verification passes were made over every entry with Claude models. Each cited page was opened where accessible, each quotation was compared with its source character by character and checked against the person or body it is attributed to (publisher, reviewer, scholar), and discrepancies were reported to the author, who decided on each. All 3,795 resulting corrections are published with the dataset as an edit log, which records for each change the row, the column, the text before and after, and, for most entries, the reason (in Russian, the working language of the project).

Throughout the verification passes, cells were brought into line with the conventions. Fields whose values could not be confirmed without close reading of the primary text are flagged in a separate column (42 entries).

Of the 1,206 unique URLs cited, 1,160 are archived in the Wayback Machine. The remaining 46, mostly retailer pages, could not be archived, but none of them are the only evidence for a work’s place in the dataset.

(3) DATASET DESCRIPTION

REPOSITORY NAME

Zenodo.

OBJECT NAME

historical_detective_fiction_dataset.xlsx, historical_detective_fiction_dataset.csv, edit_log.xlsx, edit_log.csv, README.pdf.

FORMAT NAMES AND VERSIONS

CSV, Excel, PDF.

CREATION DATES

2026-06-09–2026-09-20.

DATASET CREATORS

Marina Bulakhova (Independent Researcher, Tartu, Estonia): conceptualization, methodology, investigation, data curation, validation.

LANGUAGE

English. Titles of untranslated works and some quotations are also given in the original language and script, with literal English translations in square brackets. The *reason* field of the edit log is in Russian.

LICENSE

Creative Commons Attribution 4.0 International (CC BY 4.0).

PUBLICATION DATE

2026-09-20 (version 1); 2026-10-14 (version 2).

(4) REUSE POTENTIAL

As standalone data, the table shows which places and periods the genre takes up. Historical detective fiction set before the mid-nineteenth century is anachronistic by design: it transposes modern ideas of effective investigation into periods when justice rested on quite different principles (see, e.g., McKendry, 2019, on medieval crime fiction). The distribution of settings can therefore say much about how the past is perceived today. A scholar in memory studies might ask whether authors dare to place a just investigator in periods of state injustice, such as terror or totalitarian rule (see, e.g., Saupe, 2009, on National Socialism in crime fiction). Some settings are dated to the day: “the days of the city’s fall to Franco”, “the seven days before the promulgation of the alphabet”, or “the eve of the Exposition Universelle”, for example. Such cases may interest philosophers of history working on the concept of the historical event (see, e.g., Karla, 2021; Koselleck, 2004), as may the question of why events with nothing criminal about them, such as the eruption of Vesuvius, turn out to have detective potential.

Beyond the table itself, the dataset can serve as a bibliography of primary sources in literary studies, history and the philosophy of history, and as a bridge towards further data, such as plots, motifs, authors’ biographies and the figure of the investigator.

Limitations: the dataset is a purposive sample, not a census, and its significance columns record what was found rather than everything that exists, so they cannot be used for quantitative studies of adaptations, translations, reissues, etc. The selection may also reproduce biases of the literary record. The significance criteria reward works that have already gained visibility through film and television, prize juries, academic attention and the translation market, all of which have historically favored some groups of writers over others. The search drew on online reference sources and language models, which inherit the gaps of the web. As a result,

the dataset may underrepresent groups that are underrepresented in the literary record as a whole, such as women writers, particularly women of colour, and literatures with a small online presence. Its proportions describe the most visible part of the genre.

The author of the dataset welcomes proposals for additions, discussion of the genre attribution of particular works, and corrections of possible errors.

AI DECLARATION

Generative AI tools were used at two stages of building the dataset, as described under Method. First, large language models proposed candidate works against criteria set by the author and searched for sources on them. The models used between June and September 2026 were Claude Opus 4.8, Claude Opus 5, Claude Fable 5 and Claude Fable 5.1 (Anthropic); GPT-5.5, GPT-5.5 Pro, GPT-5.6 Sol, GPT-5.6 Sol Pro and GPT-6 Astra (OpenAI); Gemini 3.1 Pro, Gemini 3.1 Deep Think and Gemini Flash (Google); and DeepSeek-V4-Pro Preview (DeepSeek). The same models helped draft the commentary texts from sources supplied or approved by the author. The content of the columns in their final form was written by the author. Second, Claude models carried out four verification passes, comparing every quotation with its cited source and reporting discrepancies. All the decisions regarding these discrepancies were made by the author. No AI tool generated the works or sources described in the dataset. Every entry describes a real book, and the information about it is drawn from publishers' pages, reviews, award lists, catalogues and scholarly literature cited throughout the table.

Claude Fable 5.1 and Claude Opus 5.5 (Anthropic) were used to translate the author's Russian drafts of this article into English, to suggest wording and structure, and to improve overall readability.

DATA ACCESSIBILITY STATEMENT

The dataset described in this paper is openly available on Zenodo under a CC BY 4.0 licence: Version v1 https://doi.org/10.5281/zenodo.22858682, Version v2 https://doi.org/10.5281/zenodo.23373266.

ETHICS AND CONSENT

This research did not involve human participants. The dataset contains only publicly available bibliographic information about published works and their authors.

ACKNOWLEDGEMENTS

I thank my husband, the philosopher of history Andrei Gavrilin, and my former doctoral supervisor Roman Leibov for their attention to my ideas and for their valuable thoughts.

COMPETING INTERESTS

The author has no competing interests to declare.

AUTHOR CONTRIBUTIONS

Marina Bulakhova: Conceptualization, Methodology, Investigation, Data curation, Validation, Project administration, Writing – original draft, Writing – review & editing.

AUTHOR AFFILIATIONS

Marina Bulakhova orcid.org/0009-0006-0703-0415
Independent Researcher, Tartu, Estonia

REFERENCES

Browne, R. B., & **Kreiser, L. A., Jr.** (Eds.) (2000). *The detective as historian: History and art in historical crime fiction*. Bowling Green State University Popular Press.

Bulakhova, M. (2026). *Historical detective fiction: An annotated bibliographic dataset of 301 novels and series (first publications 1594–2023)* [Data set]. Zenodo. https://doi.org/10.5281/zenodo.22858681 (all versions).

Cawelti, J. G. (1976). *Adventure, mystery, and romance: Formula stories as art and popular culture*. University of Chicago Press. https://doi.org/10.7208/chicago/9780226148700.001.0001

Crime Writers' Association. (2026). *Historical Dagger*. Retrieved from https://thecwa.co.uk/awards-and-competitions/the-daggers/historical-dagger/ (last accessed: 26 September 2026).

de Groot, J. (2009). *The historical novel*. Routledge. https://doi.org/10.4324/9780203868966

Grossman, C., Spahr, J., & **Young, S.** (2022). *The index of major literary prizes in the US* [Data set]. Post45 Data Collective. https://doi.org/10.18737/CNJV1733p4520221212

Historical Writers' Association. (2025). *Crown Award terms & conditions 2025*. Retrieved from https://historicalwriters.org/awards/crown-award-terms-conditions-2025/ (last accessed: 26 September 2026).

Hubin, A. J. (2003). *Crime fiction IV: A comprehensive bibliography, 1749–2000* (Vols. 1–5). Shelburne, ON: Battered Silicon Dispatch Box.

Johnson, S. (2002). *Defining the genre: What are the rules for historical fiction?* Historical Novel Society. Retrieved from https://historicalnovelsociety.org/defining-the-genre-what-are-the-rules-for-historical-fiction/ (last accessed: 26 September 2026).

Karla, A. (2021). Controversial chronologies: The temporal demarcation of historic events. *History and Theory, 60*(1). P. 134–149. https://doi.org/10.1111/hith.12197

Koselleck, R. (2004). *Futures past: On the semantics of historical time* (K. Tribe, Trans.). Columbia University Press.

McKendry, A. (2019). *Medieval crime fiction: A critical overview*. McFarland.

Moretti, F. (2005). *Graphs, maps, trees: Abstract models for a literary history*. Verso.

Piper, A., Bamman, D., Han, C., Bjerring-Hansen, J., Long, H., Marienberg-Milikowsky, I., McEnaney, T., Orhero, M. I., Peksoy, E., Rastogi, P., Rasmussen, S., Smeets, R., Stuart, A., & **Thomsen, M. R.** (2025). Mini Worldlit: A dataset of contemporary fiction from 13 countries, nine languages, and five continents. *Journal of Open Humanities Data, 11*, 4. https://doi.org/10.5334/johd.248

Saupe, A. (2009). *Der Historiker als Detektiv – der Detektiv als Historiker: Historik, Kriminalistik und der Nationalsozialismus als Kriminalroman*. transcript. https://doi.org/10.14361/9783839411087

Todorov, T. (1977). The typology of detective fiction. In T. Todorov (Ed.), *The poetics of prose* (R. Howard, Trans., pp. 42–52). Ithaca, Cornell University Press.

Walter Scott Prize for Historical Fiction. (2025). *The Walter Scott Prize for Historical Fiction 2026: Prize rules*. Retrieved from https://www.walterscottprize.co.uk/wp-content/uploads/2025/08/WSP-Rules-2026.pdf (last accessed: 26 September 2026).

TO CITE THIS ARTICLE:
Bulakhova, M. (2026). Mapping Historical Detective Fiction Worldwide: A Bibliographic Dataset of 301 Novels and Series. *Journal of Open Humanities Data*, 12: 149, pp. 1–6. DOI: https://doi.org/10.5334/johd.676

Submitted: 26 September 2026
Accepted: 19 October 2026
Published: 20 November 2026

COPYRIGHT:
© 2026 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See https://creativecommons.org/licenses/by/4.0/.

Journal of Open Humanities Data is a peer-reviewed open access journal published by Ubiquity Press.