

## Digital appendix 1. Relationship of fields in *Trove* and the curated dataset

\* Identifies the field as part of the critical apparatus for the curated dataset

Trove Metadata Fields	Curated Dataset Fields	Description	Database table
<b>IMPORTED</b>			
article@id	chapter_id	Unique identifier for each partial or full publication of the fictional work (based on <i>Trove</i> 's article segmentation)	chapter
page	page_number	Newspaper page number where the partial or full publication of the fictional work appears	chapter
pageSequence	page_sequence	Indicates if the partial or full publication of the fictional work appears as part of a supplement or in a special section	chapter
identifier	chapter_url	Persistent URL for the digitized version of the partial or full publication of the fictional work in <i>Trove</i>	chapter
trovePageUrl	page_url	Persistent URL for the digitised page where the instalment or full publication of a fictional work appears	chapter
illustrated	illustrated	Indicates whether the partial or full publication of the fictional work is illustrated (Y/N)	chapter
correctionCount	corrections	Number of corrections made to <i>Trove</i> 's OCR text of a partial or full publication of the fictional work (each correction may involve multiple lines)	chapter
lastCorrection	last_corrected	Date of most recent crowd-sourced correction to <i>Trove</i> 's OCR text of a partial or full publication of the fictional work	chapter
wordCount	word_count	Number of words in a partial or full publication of the fictional work	chapter
articleText	article_text	Full text of the partial or full publication of the fictional work, incorporating any corrections to the OCR text in <i>Trove</i>	chapter
title@id	newspaper_table_id	Unique identifier for each digitized newspaper in <i>Trove</i>	newspaper & title
state	newspaper_	State or colony in which the	newspaper

	location	newspaper was predominantly published	
startDate	start_date	Date of the earliest issue of the newspaper digitized by <i>Trove</i>	newspaper
endDate	end_date	Date of the most recent issue of the newspaper digitized by <i>Trove</i>	newspaper
issn	issn	Newspaper's International Standard Serial Number	newspaper
<b>EXPANDED</b>			
heading	chapter_number	Chapter number, if any, for the partial or full publication of the fictional work	chapter
	chapter_title	Chapter title, if any, for the partial or full publication of the fictional work	chapter
	publication_title	Title of the fictional work as it appears in the publication event	title
	publication_author	Attribution of the fictional work as it appears in the publication event (including "Unattributed")	title
	signature	Other works by the same author listed in the publication event	title
	publication_source	Source of the fictional work as identified in the publication event	title
	inscribed_gender	Gender of the author indicated in the publication event (including "Uninscribed")	title
	inscribed_nationality	Nationality of the author indicated in the publication event (including "Uninscribed")	title
	name_category	Form in which the author's name is represented in newspaper	title
	additional_info	Other information about the author or fictional work supplied in the publication event (e.g. copyright details; translation details; whether a title is part of a series)	title
title	newspaper_title	Name of the newspaper as listed in <i>Trove</i>	newspaper
	newspaper_details	Information about the newspaper provided as part of <i>Trove</i> 's title information	newspaper
date	day	Day of the week the partial or full publication of the fictional work appeared	chapter
	date	Date of the month the partial or full publication of the fictional work appeared	chapter
	month	Month in which the partial or full	chapter

		publication of the fictional work appeared	
	year	Year in which the partial or full publication of the fictional work appeared	chapter
	full_date	Date (YYYY-MM-DD) of the partial or full publication of the fictional work	Chapter
<b>CREATED</b>			
N/A	title_id	Unique identifier for the fictional work in a particular newspaper at a particular time (i.e. the unique identifier of a publication event)	title
	common_title	Title of the fictional work as is most commonly represented throughout the database and/or in catalogues, bibliographies and reference sources	title
	length	Indicates whether a publication event is extended – either by serialization (0) or by comprising more than 10,000 words in a single newspaper issue (8) – or concluded in a single newspaper issue (1)	title
	span_start	Date (YYYY-MM-DD) of the first identified instalment of the fictional work	title
	span_end	Date (YYYY-MM-DD) of the final identified instalment of the fictional work	title
	curated_dataset	Indicates whether a publication event is part of the curated dataset analysed in <i>A World of Fiction</i> (Y/N). Editing of bibliographical data is prevented for records listed as “Y”	title
	* information_source	Lists sources of information about the fictional work, including its publication or attribution	title
	found_source	Titles of other periodicals or databases where the fictional work is also published (not comprehensive)	title
	author_id	A unique identifier for each author name in the curated dataset (including “Unknown”)	title & author
	common_name	Author of a work as that individual is most commonly represented throughout the database and/or in catalogues, bibliographies and	author

		reference sources (if discoverable, otherwise listed as “Unknown”)	
	other_name	Other names used by or ascribed to the author (to improve discoverability)	author
	gender	Author’s gender, if discoverable (otherwise “Unknown”)	author
	nationality	Author’s nationality, if discoverable (otherwise “Unknown”)	author
	* nationality_details	Other information about nationality, including where authors were born, or spent considerable time, in countries other than their designated “nationality”	author
	* author_details	Lists reference sources where information about authorship was discovered	author
	newspaper_type	Categorises newspapers based on whether they were published within ten kilometres (“metropolitan”), between ten and thirty kilometres (“suburban”), or more than thirty kilometres (“provincial”) from the state or colony’s central business district	newspaper
	newspaper_common_title	A single title to align as one publication different banners of the same newspaper	newspaper
	start_year	Year of publication for the earliest issue of the newspaper digitized by <i>Trove</i>	newspaper
	end_year	Year of publication for the most recent issue of the newspaper digitized by <i>Trove</i>	newspaper