Voices from the Past: A Digital Exploration of Australian World War I Diaries

Ashley Dennis-Henderson

December 4, 2025



THE UNIVERSITY of ADELAIDE

Acknowledgements

I would like to thank:

- The State Library of New South Wales for providing the transcriptions of the diaries used in this project.
- The AIF Project for providing service records.
- My supervisors Matthew Roughan, Jono Tuke, and Gareth Pritchard for their guidance in this project.
- The aaDH for providing a bursary to help me attend this conference.

Talk Overview

1. Project Motivation
2. Data
3. Who were the Diarists?
4. What did they Write About?
5. Where to from here?

Project Motivation

PhD Aims

- To investigate how distant reading techniques can be used to aid historians in analysing large collections of historic text.
- To analyse Australian World War I diaries using distant reading techniques in order to understand what the soldiers wrote about and how they felt over the course of the war.

Close and Distant Reading

Close reading is the careful reading and interpretation of a text to identify central themes and their development.

Distant reading is the use of computational techniques to create an abstract view of the text and hence visualise global features and trends.

Data

State Library of New South Wales Collection

- Began as the European War Collecting Project.
- Documents have been digitised and transcribed.
- 504 cleaned diaries are ready to be analysed.
- I am still in the process of cleaning the letters in this collection.

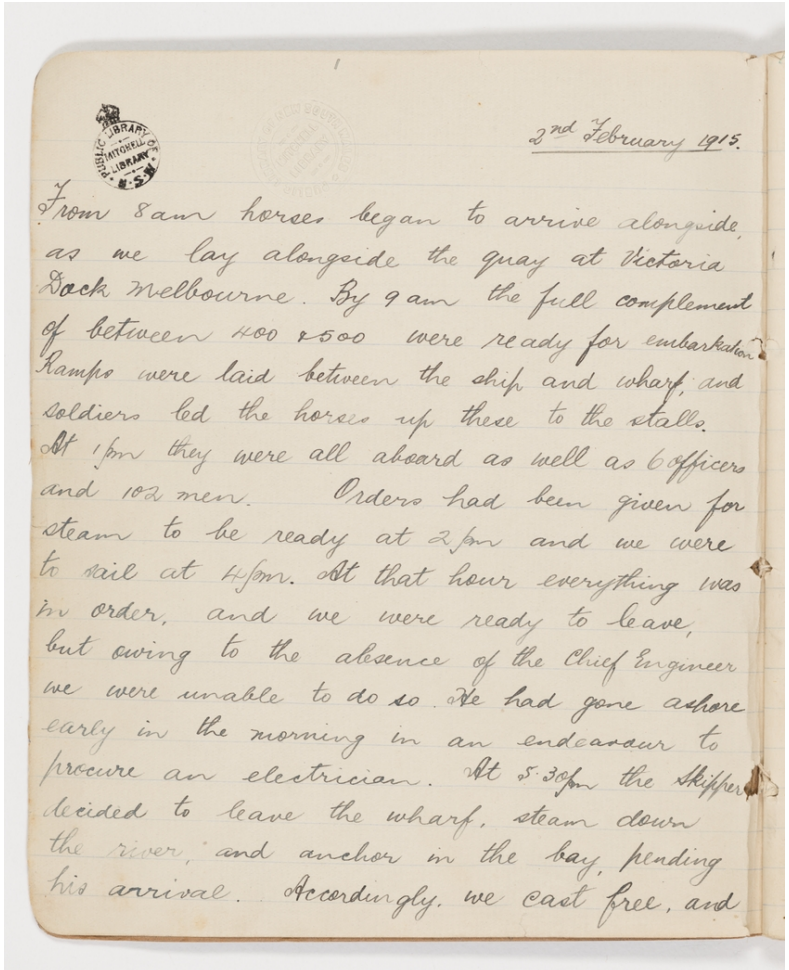


Figure: Gilroy, Norman. *Norman Thomas Gilroy war diary, Volume 1: 2 February-29 April 1915*. Mitchell Library, State Library of New South Wales.

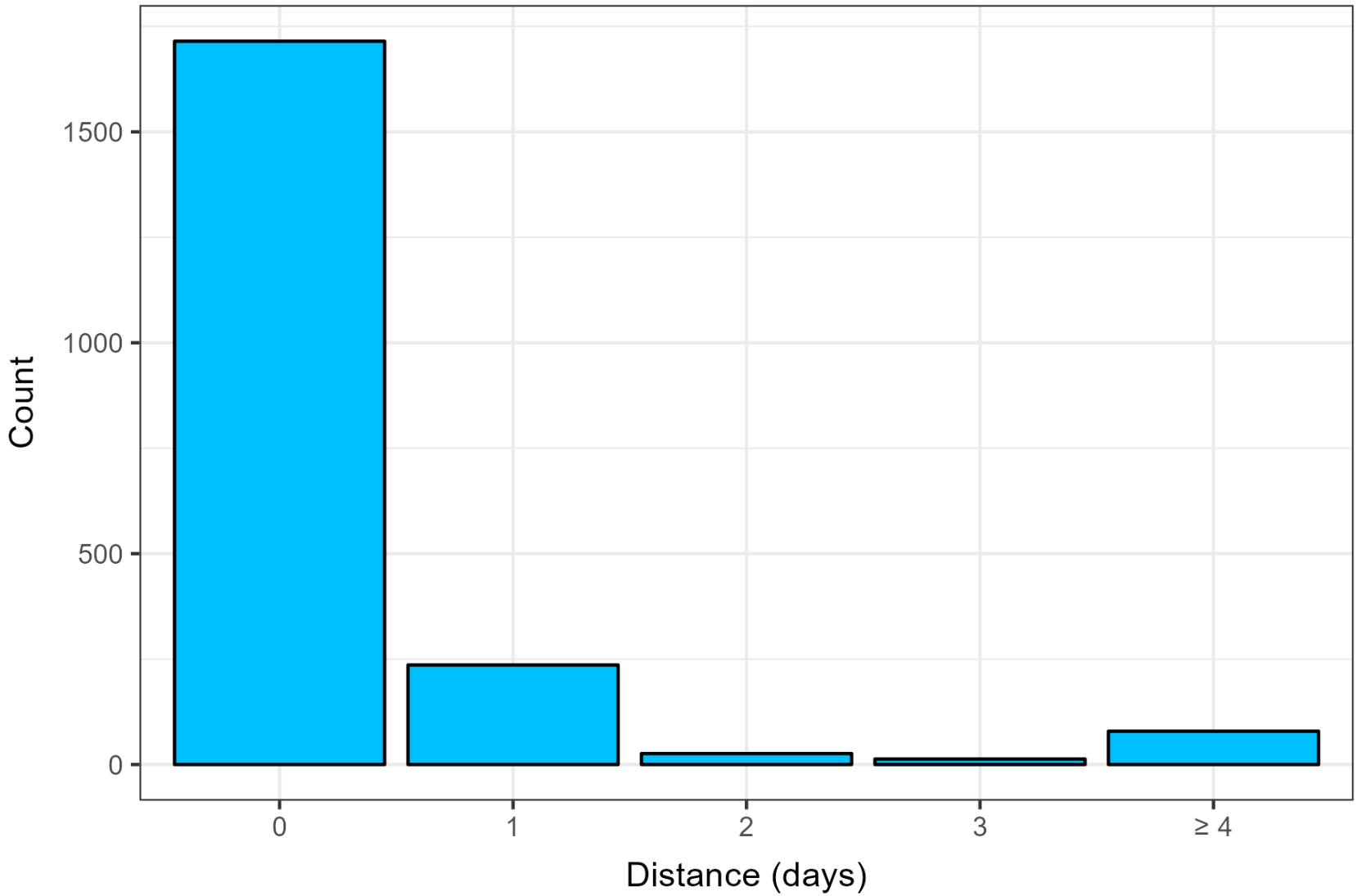
AIF Project

- Began in 1987 at the Australian Defence Force Academy (ADFA).
- Aimed to create a comprehensive computer database of Australian Imperial Force (AIF) service records for WWI.
- https://www.aif.adfa.edu.au/index.html

Data Cleaning

- Create complete and accurate metadata.
- Remove diaries from those not enlisted in the AIF, nurses, and honorary enlistees.
- Remove transcribers notes and pre-/post- text.
- Extract dates and remove out-of-order diaries.

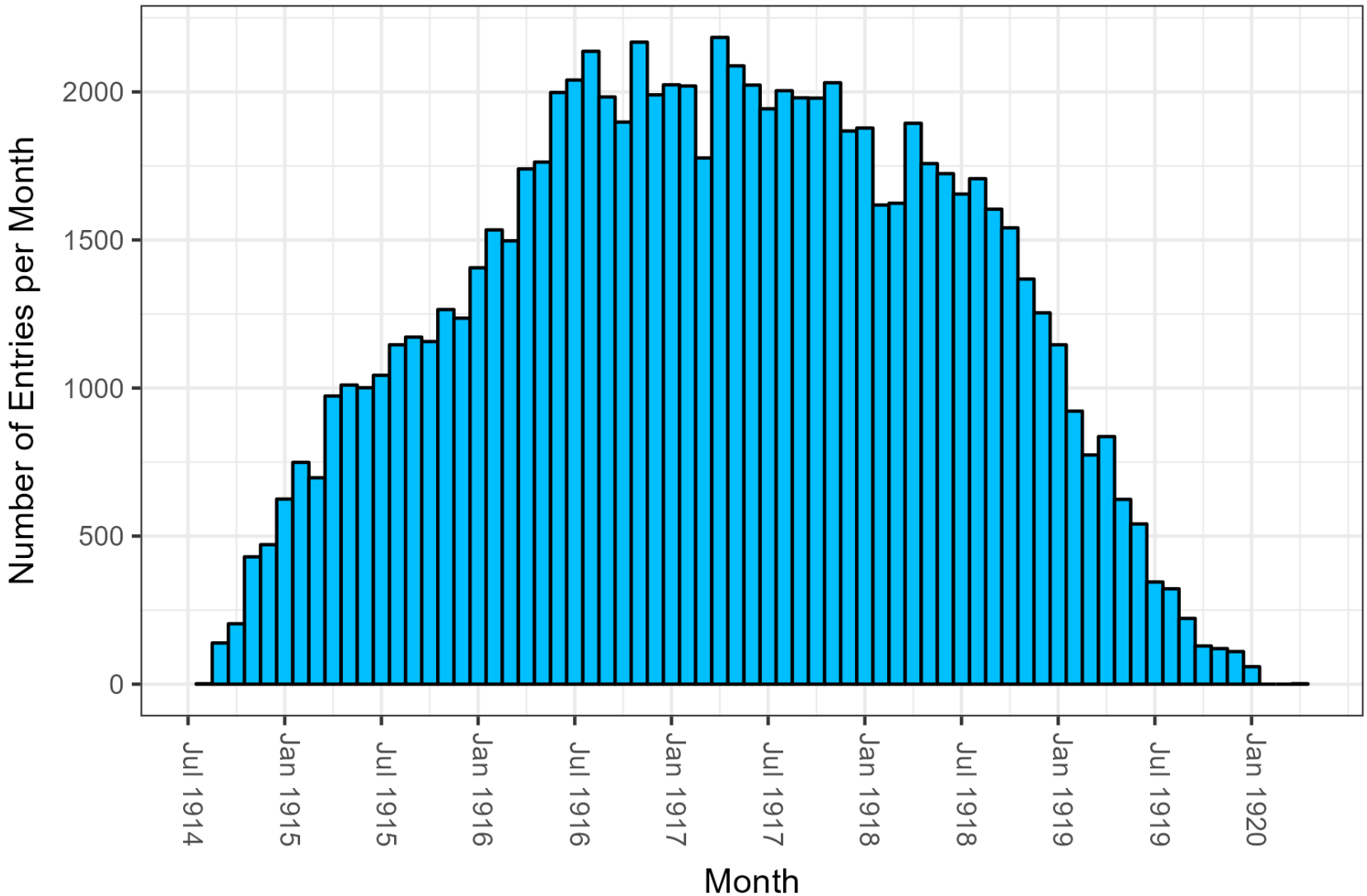
Date Extraction Accuracy



Diary Metrics

- 504 diaries
- 195 authors
- 85,578 entries
- Approximately 7 million words

Number of Entries



Who were the Diarists?

Address at Enlistment

	Count	%
NSW	139	71.3
VIC	15	7.7
QLD	11	5.6
WA	4	2.1
SA	2	1.0
TAS	3	1.5
England	2	1.0
Unknown	19	9.7

Table 1: Diarists' state of residence.

Address at Enlistment

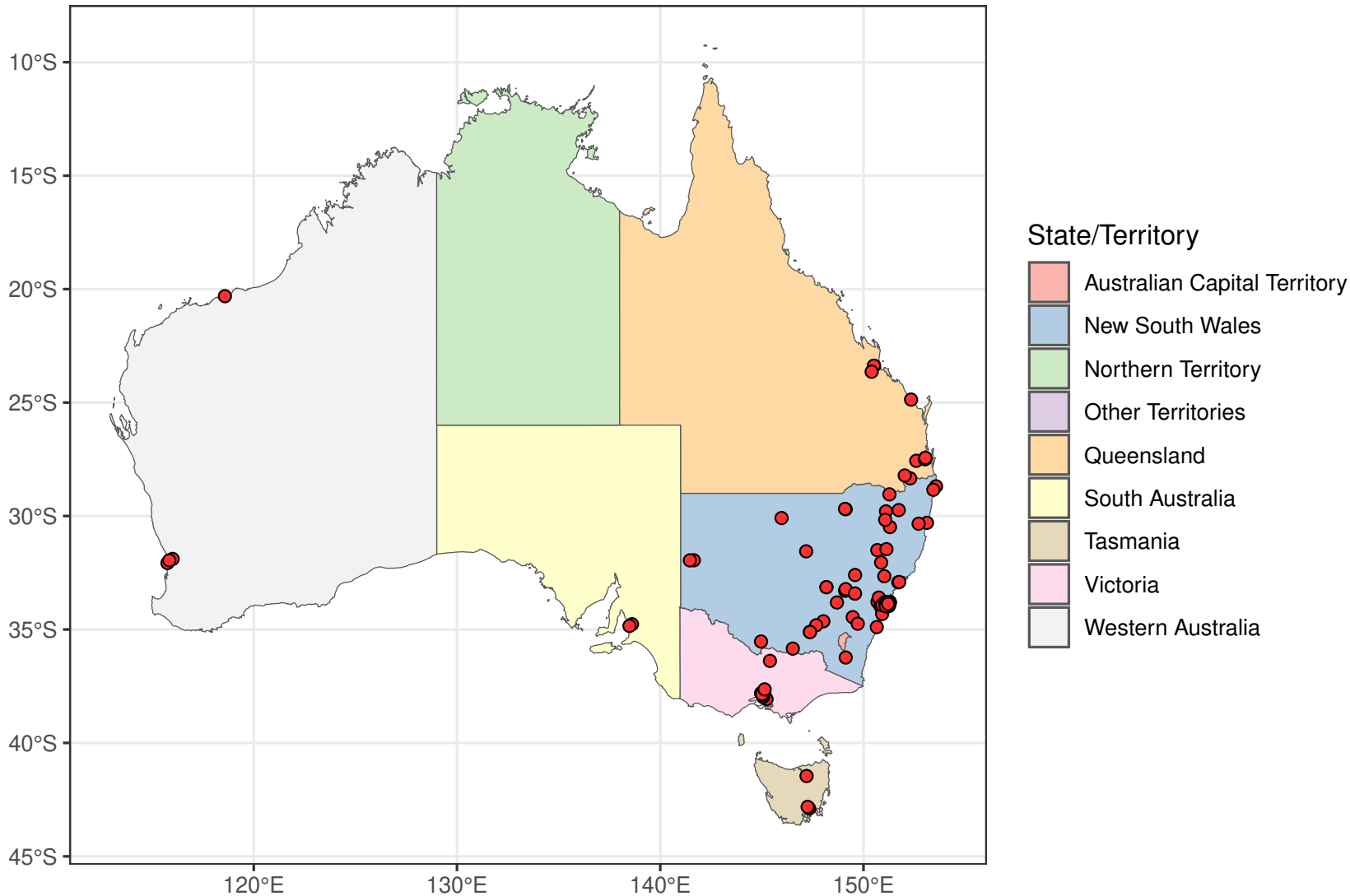
	Count	%
NSW	139	71.3
VIC	15	7.7
QLD	11	5.6
WA	4	2.1
SA	2	1.0
TAS	3	1.5
England	2	1.0
Unknown	19	9.7

Table 1: Diarists' state of residence.

	Count	%
NSW	164,030	39.4
VIC	112,399	27.0
QLD	57,705	13.8
WA	32,231	7.7
SA	34,959	8.4
TAS	15,485	3.7

Table 2: Official enlistment figures per state.

Address at Enlistment



Enlistment Location

	Count	%
NSW	134	68.7
VIC	13	6.7
QLD	11	5.6
WA	4	2.1
SA	3	1.5
TAS	2	1.0
Unknown	28	14.4

Table 3: Diarists' enlistment states.

Enlistment Location

	Count	%
NSW	134	68.7
VIC	13	6.7
QLD	11	5.6
WA	4	2.1
SA	3	1.5
TAS	2	1.0
Unknown	28	14.4

Table 3: Diarists' enlistment states.

	Count	%
NSW	164,030	39.4
VIC	112,399	27.0
QLD	57,705	13.8
WA	32,231	7.7
SA	34,959	8.4
TAS	15,485	3.7

Table 4: Official enlistment figures per state.

Why More Diarists from NSW?

Why More Diarists from NSW?

- Were soldiers from NSW more likely to offer their diary as it was going to their state's library?

Why More Diarists from NSW?

- Were soldiers from NSW more likely to offer their diary as it was going to their state's library?
- Were the representatives and newspaper advertisements targeted more at soldiers from NSW?

Why More Diarists from NSW?

- Were soldiers from NSW more likely to offer their diary as it was going to their state's library?
- Were the representatives and newspaper advertisements targeted more at soldiers from NSW?
- Were library staff more likely to accept a document if it was written by someone from NSW?

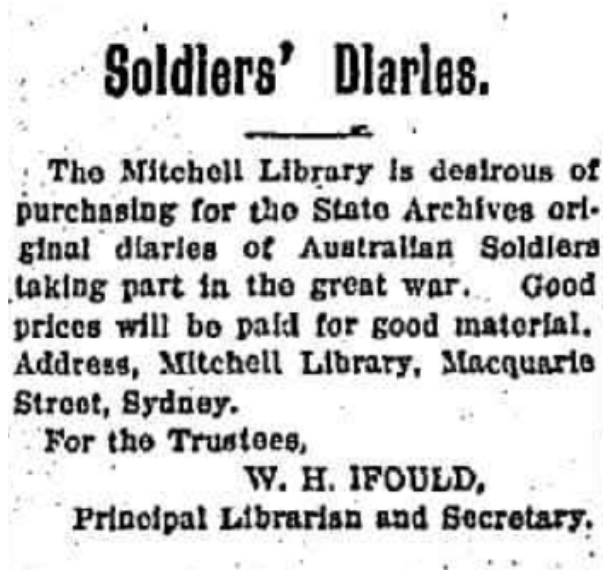
Newspaper Advertisements



SOLDIERS' DIARIES.

Thousands of volumes have already been published on the great war, and for some years to come thousands more will be issued. Some will be comprehensive, and others will deal with special phases or with the achievements of selected forces. When the part played by the Australians is written the official despatches will furnish much useful data, but for the personal feelings, doings, and relationships of the men, their thoughts and actions, the diaries or journals kept by the men themselves will be of the greatest value. The Trustees of the Mitchell Library recognise the importance of collecting and preserving these records, and by an advertisement in our columns announce their willingness to buy such manuscripts. Originals are wanted, not copies. We advise anyone having in their possesion such diaries to make copies for their own use and offer the originals to the Mitchell Library, which is the treasure house of Australia for choice manuscripts and rarest books, and is the best depositary for the diaries and important letters written by the men who gave Australia a place on the scroll of Fame.

(a) Daily Observer (NSW), 14 December 1918.



Soldiers' Diaries.

The Mitchell Library is desirous of purchasing for the State Archives original diaries of Australian Soldiers taking part in the great war. Good prices will be paid for good material. Address, Mitchell Library, Macquarie Street, Sydney.

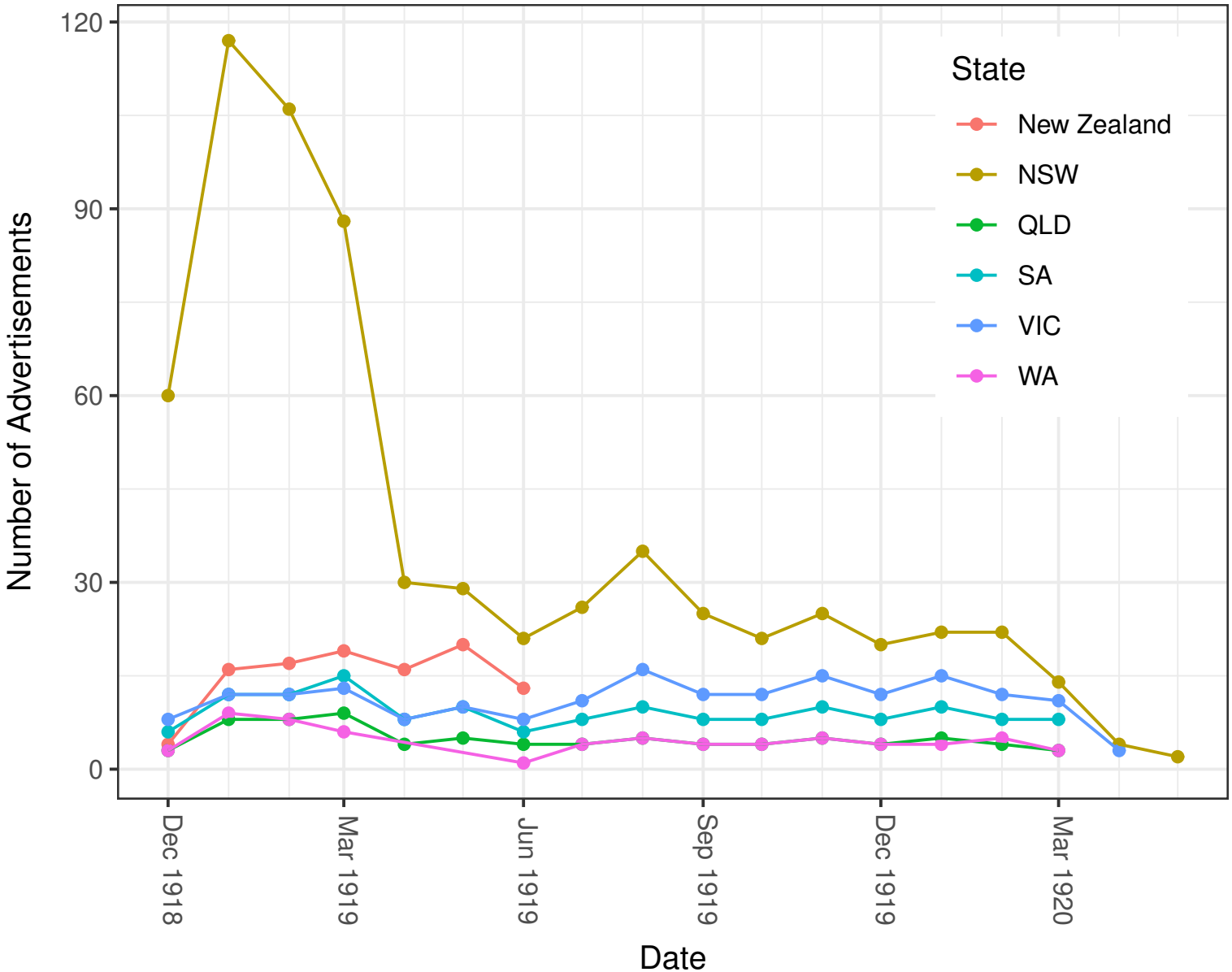
For the Trustees,

W. H. IFOULD,

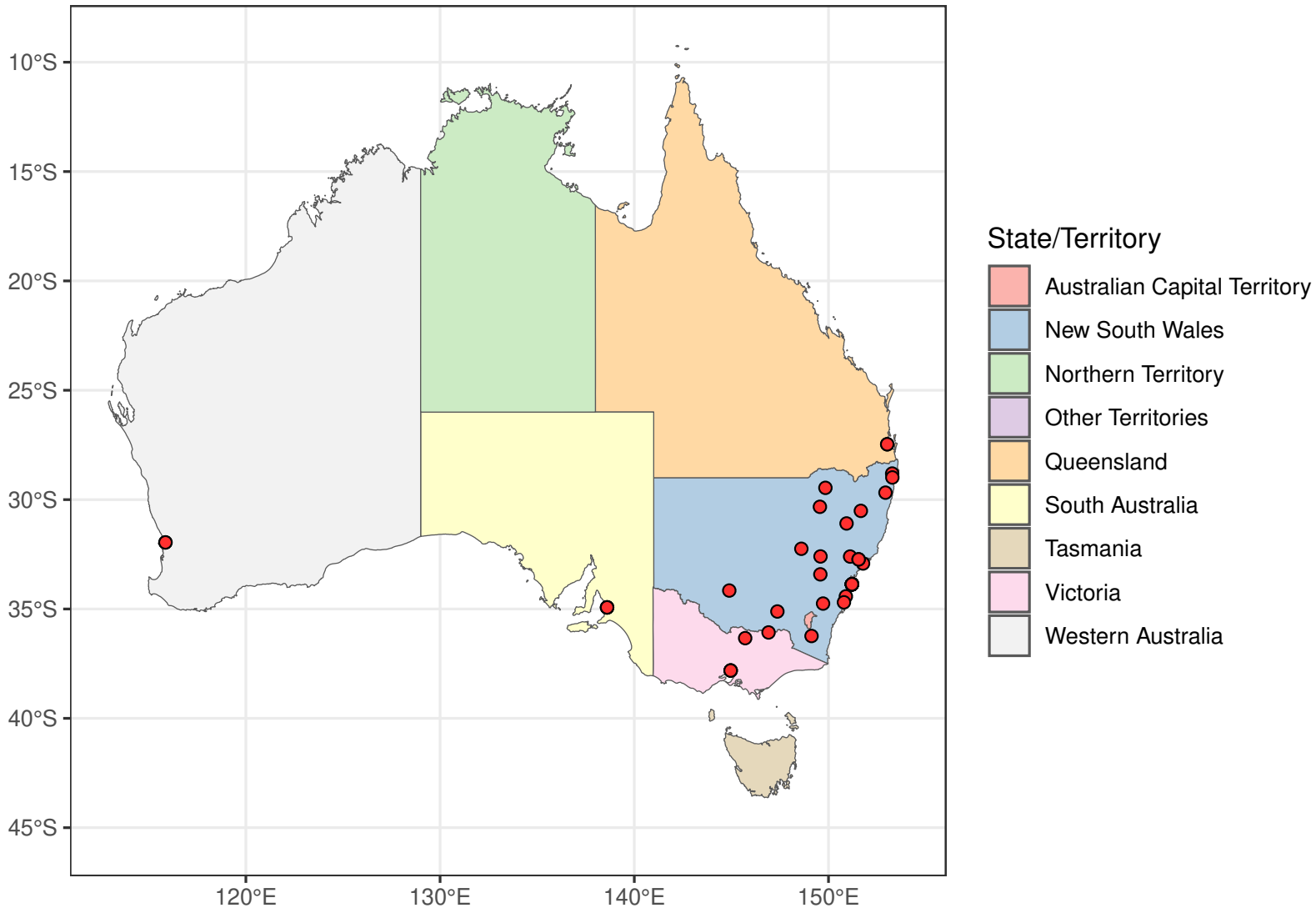
Principal Librarian and Secretary.

(b) The North Western Courier (NSW), 12 December 1918.

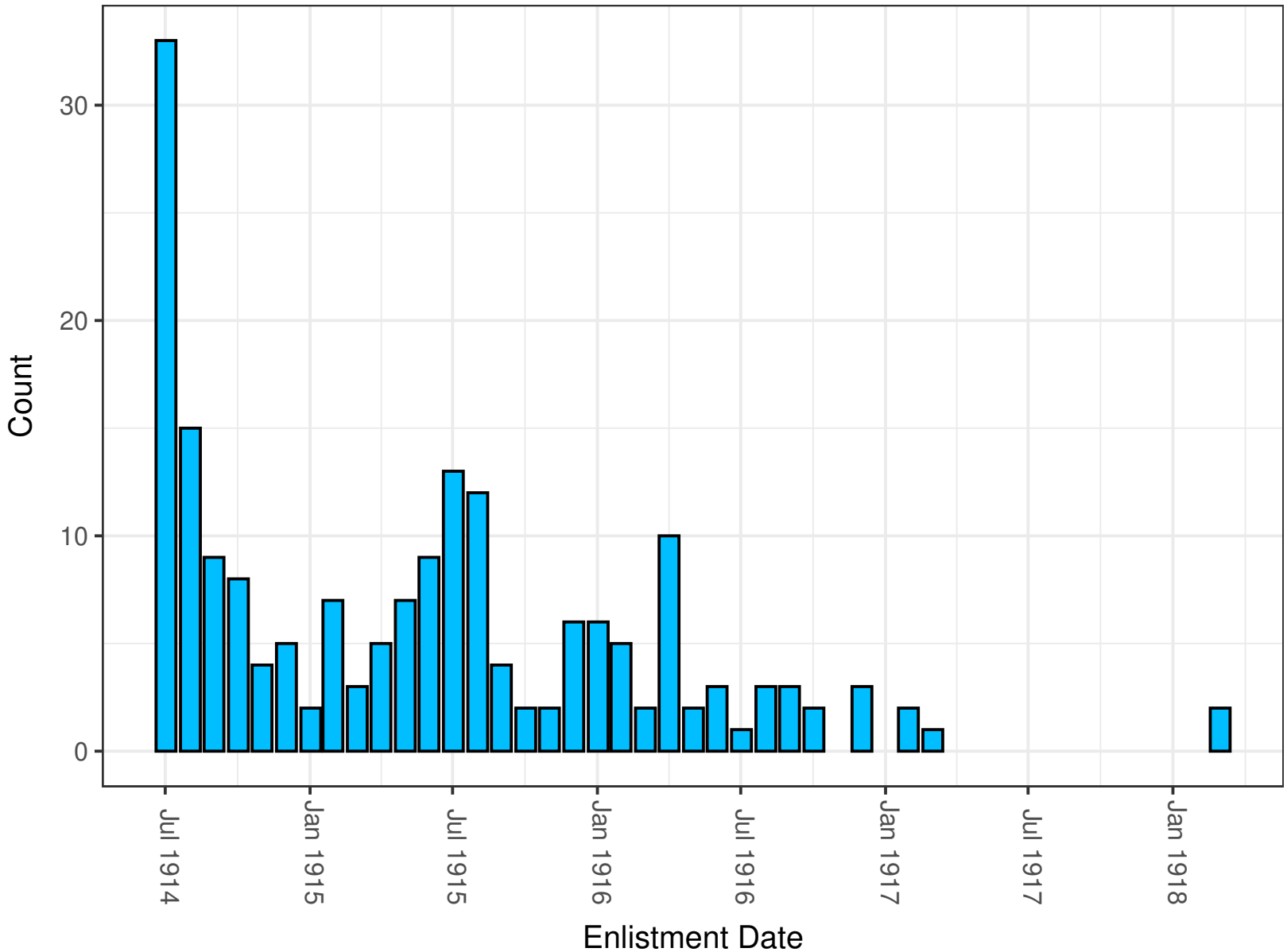
Total Advertisements



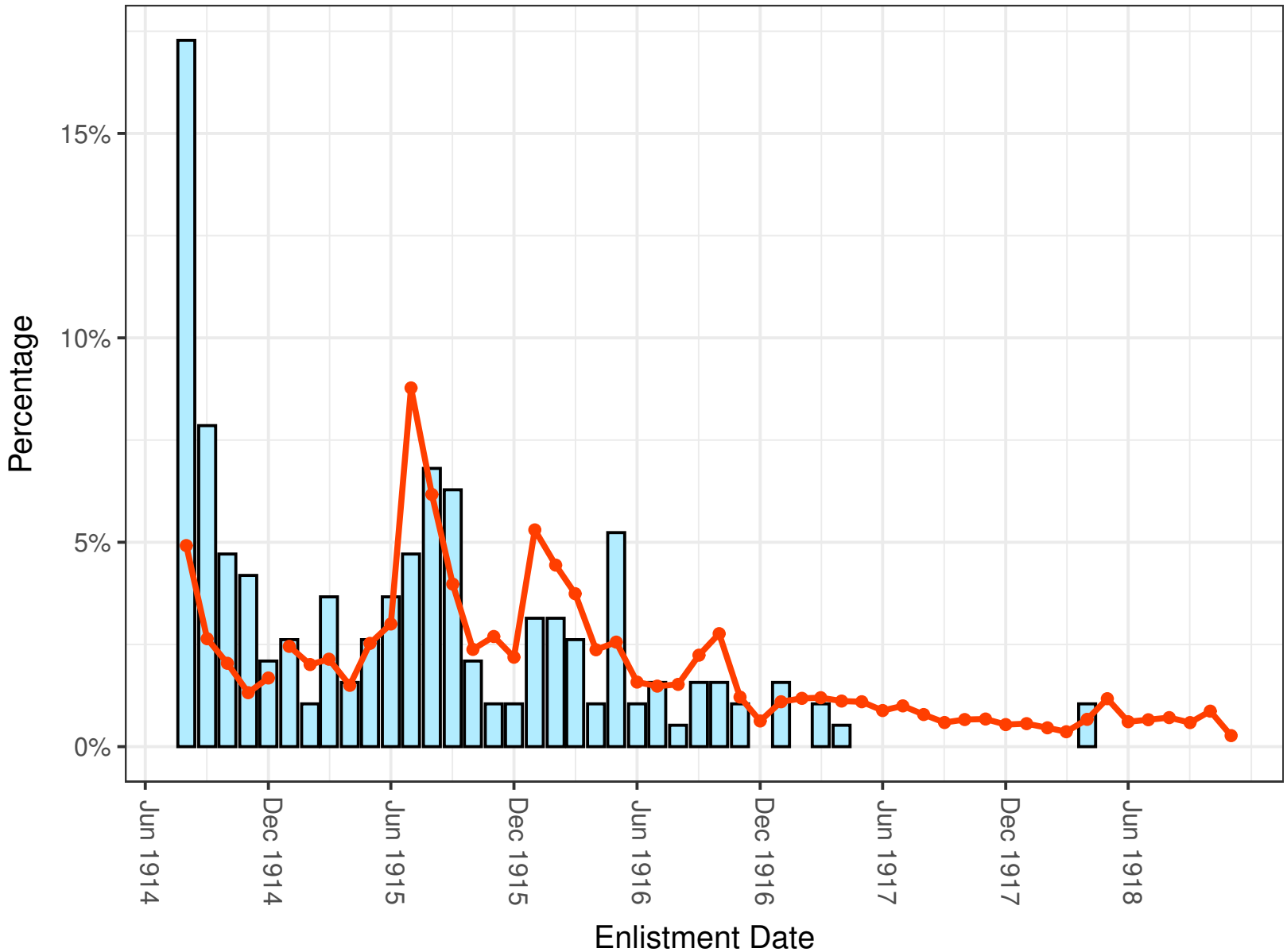
Newspaper Map



Enlistment Date



Enlistment Date



Fate

	Count	%
Returned	164	84.1
Died	14	7.2
Abroad	5	2.6
Discharged	2	1.0
Unknown	10	5.1

Table 5: Fate of our diarists.

Fate

	Count	%
Returned	164	84.1
Died	14	7.2
Abroad	5	2.6
Discharged	2	1.0
Unknown	10	5.1

Table 5: Fate of our diarists.

Official Statistics:

- 14.5% of those who enlisted died
- 18.2% of those who embarked died

Enlistment Rank

	Count	%
Other Ranks	148	75.9
Commissioned	26	13.3
NCOs	17	8.7
Unknown	4	2.1

Table 6: Diarists' enlistment rank.

Enlistment Rank

	Count	%
Other Ranks	148	75.9
Commissioned	26	13.3
NCOs	17	8.7
Unknown	4	2.1

Table 6: Diarists' enlistment rank.

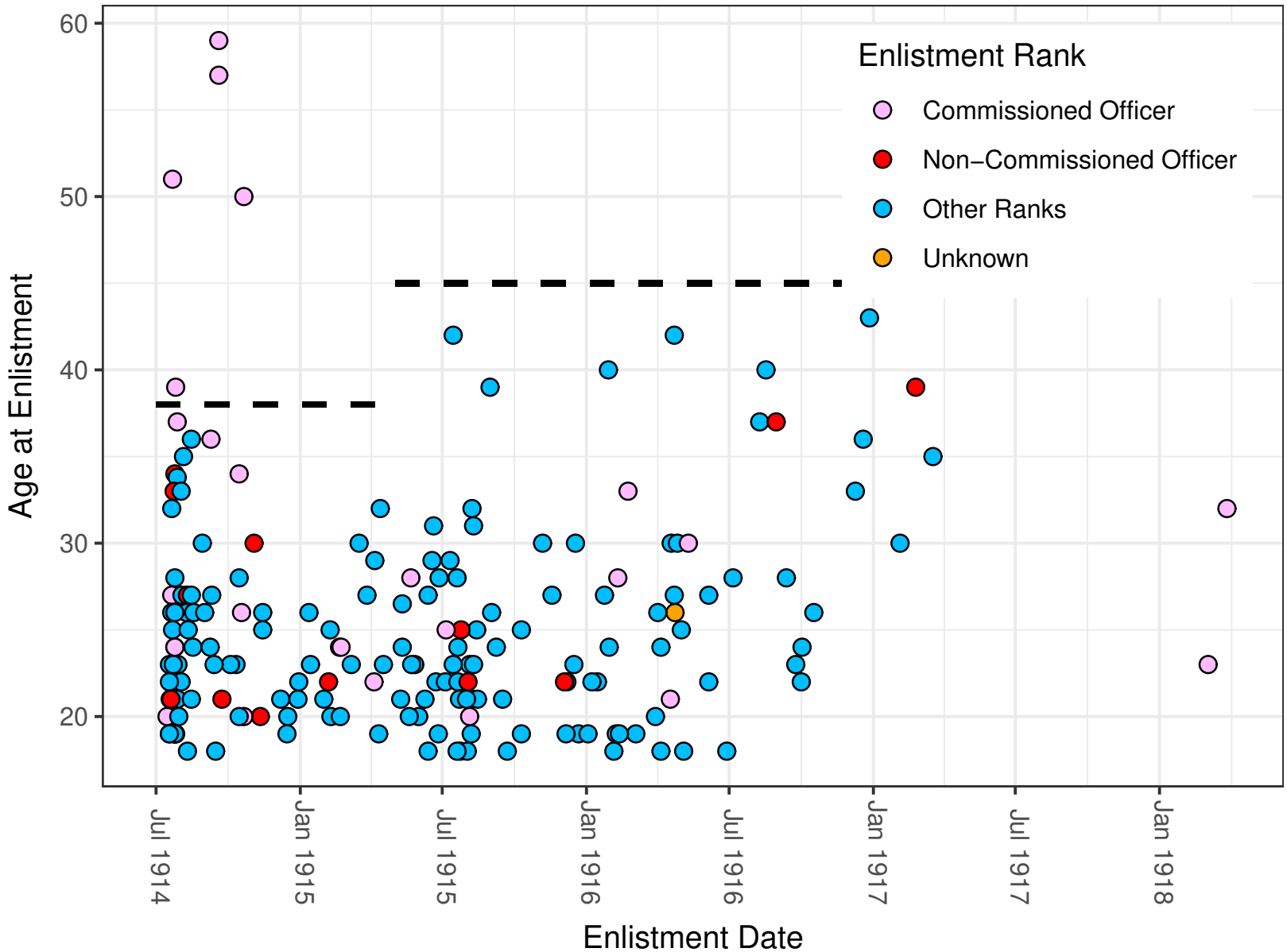
Robson's Sample:

- 78% other ranks
- 5% commissioned officers
- 17% non-commissioned officers (NCOs)

Occupations Prior to Enlistment

	Diarists		Official	
Category	**Count**	**%**	**Count**	**%**
Clerical	41	21.0	24,538	7.3
Country Callings	33	16.9	57,535	17.2
Labourers	23	11.8	99,916	29.9
Miscellaneous	14	7.2	14,229	4.3
Professional	40	20.5	15,879	4.7
Seafaring	1	0.5	6,645	2.0
Tradesmen	42	21.5	113,775	34.0
Nurses	0	0.0	2,074	0.6
Unknown	1	0.5	0	0.0

Age at Enlistment



Our data is biased . . .

Our data is biased . . .

. . . so what do we do now?

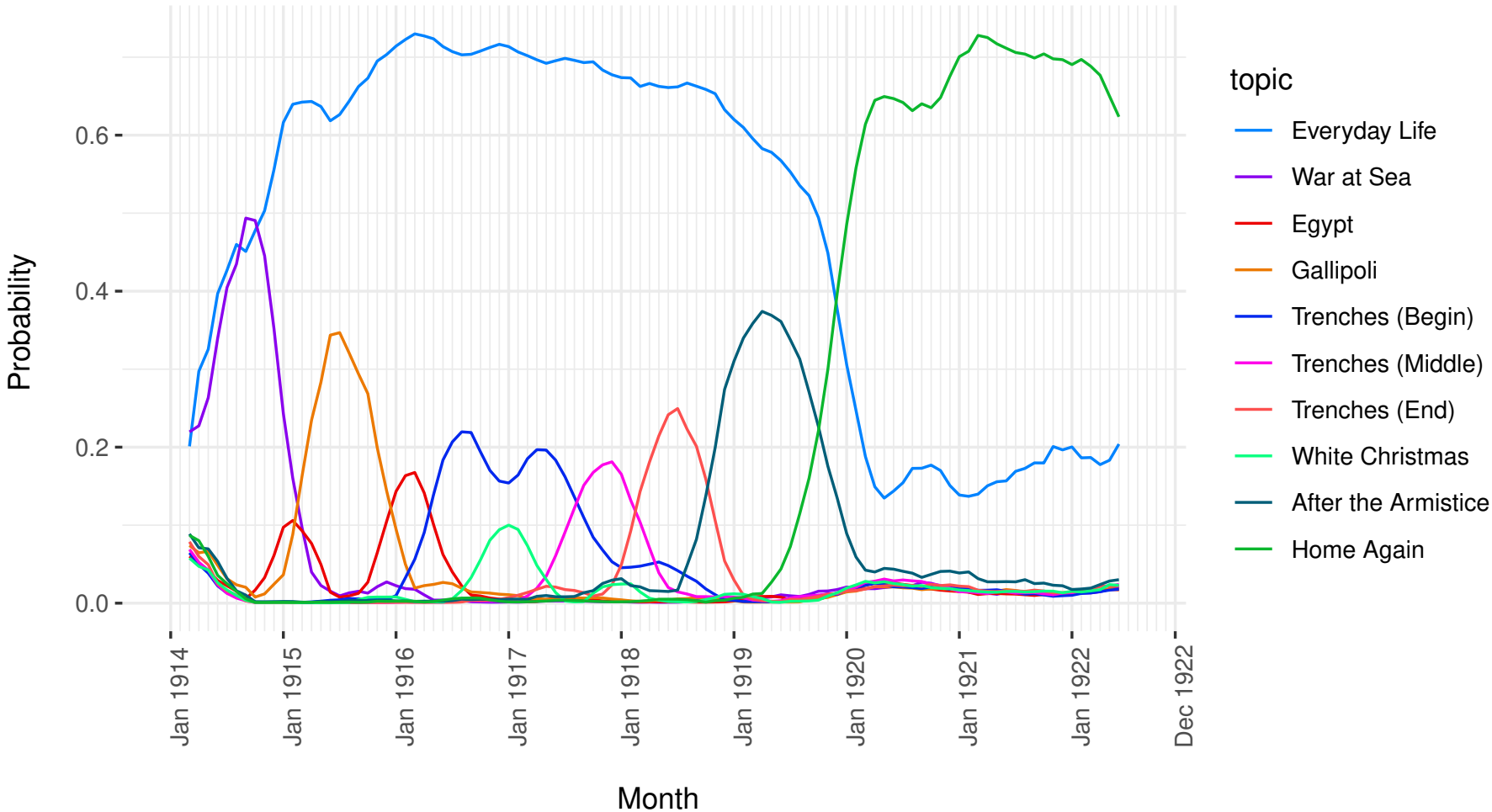
What did they Write About?

Word Frequencies

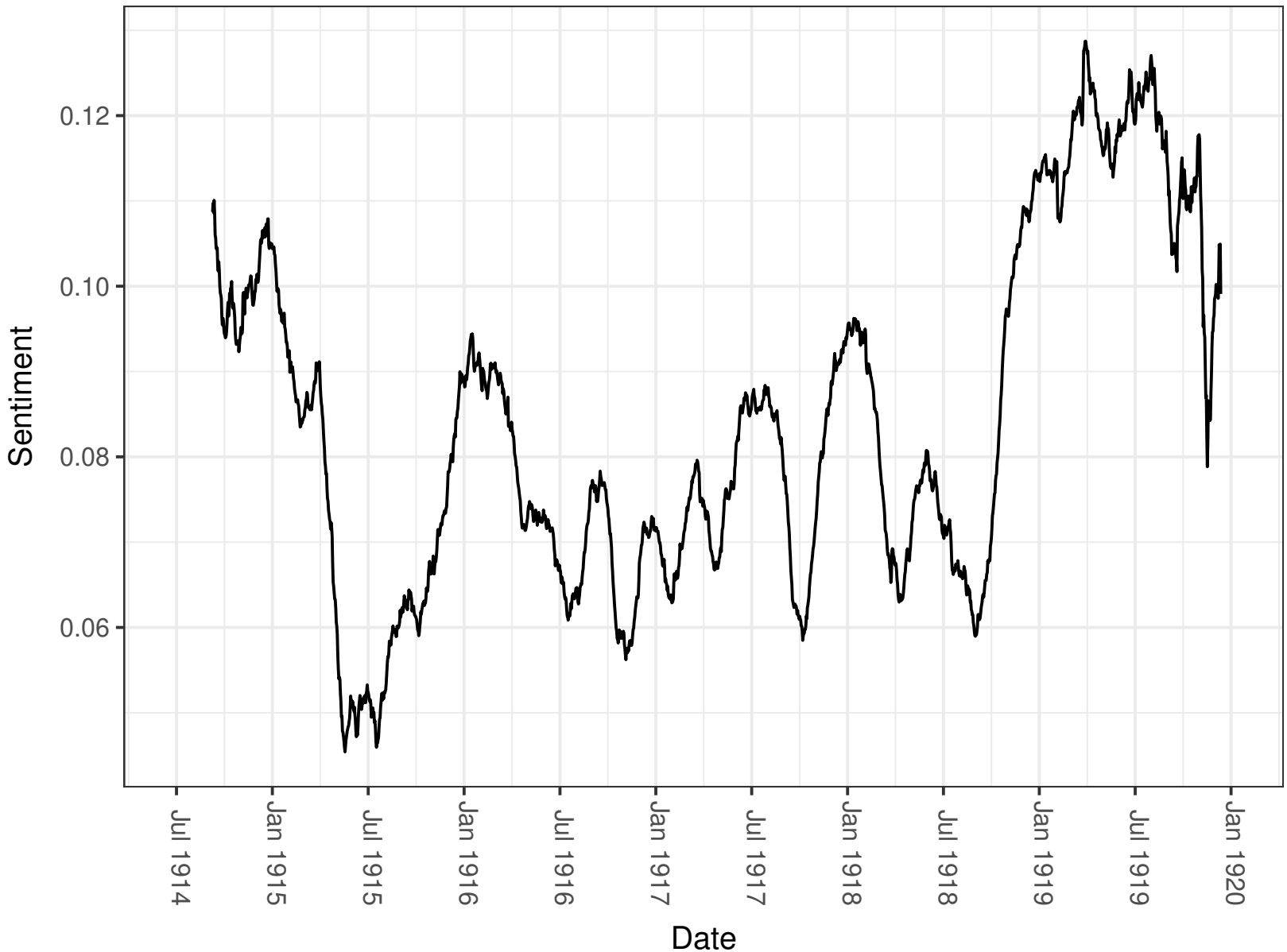


Topic Modelling

Topic Modelling



Sentiment Analysis



Where to from here?

Future Work

- Investigating other methods for topic modelling and sentiment analysis.
- Mapping soldiers' journeys through location extraction.
- Investigating the affects of censorship on letters.
- Statistical analysis of all AIF service records with comparisons to 1911 census data.
- Investigating the use of ERGMs on the text.

Questions?

https://www.ashleydennishenderson.com/

ashley.dennis-henderson@adelaide.edu.au