



ADM+S WORKING PAPER SERIES

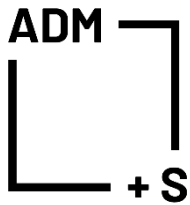
> Measuring Intermediary News Diversity Google News in Australia

October 2023

Authors:

James Meese, Associate Investigator
Abdul Karim Obeid, Data Engineer
Daniel Angus, Chief Investigator
Axel Bruns, Chief Investigator

WORKING PAPER 007



Acknowledgement of Country

In the spirit of reconciliation, we acknowledge the Traditional Custodians of country throughout Australia and their connections to land, sea and community. We pay our respect to Elders past and present and extend that respect to all Aboriginal and Torres Strait Islander peoples today.

Suggested citation. Meese, J, et al. (2023)

Measuring Intermediary News Diversity: Google News in Australia, ADM+S Working Paper Series 2023 (7), ARC Centre of Excellence for Automated Decision-Making and Society, DOI: 10.25916/xk6y-a642.

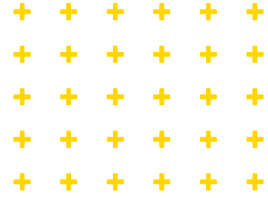
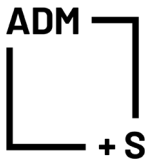
Copyright © 2023 James Meese, Abdul Obeid, Daniel Angus, Axel Bruns. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC BY 4.0). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited.

ARC Acknowledgement

This research was conducted by the ARC Centre of Excellence for Automated Decision-Making and Society (CE200100005), and funded by the Australian Government through the Australian Research Council



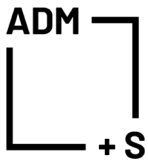
Australian Government
Australian Research Council



ABSTRACT

This Working Paper presents results from a sub-project established under the banner of the ADM+S Australian Search Experience 'citizen science' research project. Our aim was to understand what sort of news is presented to Australians through Google News. The project investigated 20 keywords related to Australian politics and global issues, including names of politicians, political parties, global sporting events and various COVID-related terms. These keywords were organised into two keyword sets: 'Australian Politics' and 'Global Issues'. We found that Google News search results were not personalised with respect to demographics or geography. However, an isolated subset of domains frequently occupied the first page of the results, and results were not always localised to Australia. These findings help us better understand how search engines present news content to ordinary internet users.

KEYWORDS: GOOGLE NEWS, JOURNALISM, MEDIA DIVERSITY, EXPOSURE DIVERSITY, RECOMMENDATION, PLATFORMS.



AUTHORS

James Meese is a senior lecturer at RMIT University and an associate investigator with the ARC Centre of Excellence for Automated Decision-Making and Society (ADM+S). His research interests include media and telecommunications policy, journalism and intellectual property. James has received research funding from the Australian Research Council, Meta, the International Association of Privacy Professionals and the Australian Communications Consumer Action Network. He regularly publishes work in leading media and communication journals, and his most recent book is *Digital Platforms and the Press* (2023).

Abdul Karim Obeid is a researcher and data engineer at the Queensland University of Technology (QUT) node of the ADM+S. Among various computational topics relating to computer science and information theory, Abdul specialises in the acquisition, transformation, storage and analysis of data. His most recent works have examined popular social networking sites and search engines.

Daniel Angus FQA is Professor of Digital Communication, Director of the QUT Digital Media Research Centre, and a chief investigator and Chair of Infrastructure with the ADM+S. His research examines issues at the intersection of technology and society, with a focus on algorithms, misinformation and new methods to study digital society.

Axel Bruns is an Australian Laureate Fellow and a professor in the Digital Media Research Centre at Queensland University of Technology in Brisbane, Australia, and a chief investigator at the ADM+S. His books include *Are Filter Bubbles Real?* (2019) and *Gatewatching and News Curation: Journalism, Social Media, and the Public Sphere* (2018). His current research focuses on the study of public communication in digital and social media environments, with a particular focus on the dynamics of polarisation, partisanship and problematic information, and their implications in understanding the contemporary public sphere; his work draws especially on innovative new methods for analysing 'big social data'. He served as President of the Association of Internet Researchers in 2017–19.

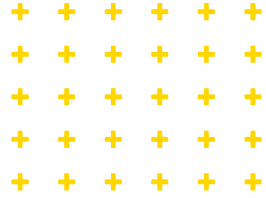
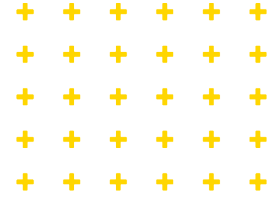


TABLE OF CONTENTS

1.	Executive Summary	6
2.	Background	7
3.	Method	8
4.	Research Approach	9
5.	Results	10
	Personalisation	10
	Frequency	11
	Location	13
	Limitations	15
6.	Appendix: Additional Details On Method	16
7.	References	16



1. EXECUTIVE SUMMARY

This Working Paper presents results from a sub-project established under the banner of the ADM+S Australian Search Experience ‘citizen science’ research project. Our aim was to understand what sort of news is presented to Australians through Google News.¹ This project specifically examined the separate Google News website, (<https://news.google.com/home>), not organic search results or the Top News carousel. We investigated 20 keywords related to Australian politics and global issues, including names of politicians, political parties, global sporting events and various COVID-related terms. These keywords were organised into two keyword sets: ‘Australian Politics’ and ‘Global Issues’.

We found the following:

- Google News search results were not personalised with respect to demographics or geography.
- An isolated subset of domains frequently occupied the first page of the Google News search results.
- The Australian Broadcasting Corporation was the only Australian news domain that appeared in each keyword set.
 - The most observed websites in the ‘Australian Politics’ keyword set were the Guardian Australia, the ABC, the Sydney Morning Herald, the Daily Mail (Australia/United Kingdom), news.com.au, Sky News and The New Zealand Herald.
 - The most observed outlets in the ‘Global Issues’ keyword set were the Guardian (U.K.), the ABC, renew.biz, The New York Times, CNN, Environmental Leader and Reuters.
- The ‘Australian Politics’ search terms were not always localised.
- The most frequently observed domains for the ‘Australian Politics’ keyword set were Australian.
- Only one Australian source was frequently observed in the ‘Global Issues’ keyword set.
- There was a preference for domains from the United States and the United Kingdom in the ‘Global Issues’ keyword set.

¹ The project focuses on Google News, which is different from Google News Showcase, a product that features selected outlets that have entered into financial agreements with Google.



2. BACKGROUND

Intermediaries have become increasingly responsible for the distribution of online news content. This means that recommender systems and search engines are now functioning as central intermediaries between news outlets and citizens. There has also been a growing focus on the ‘demand side’ of media diversity, otherwise known as exposure diversity (Napoli, 2011; Helberger, 2019). This involves assessing the diversity of news sources that people are receiving and accessing.

Recent work has debunked some of the more extreme concerns in this area, namely the societal effect of filter bubbles and echo chambers (Bruns, 2022). We know that people are more likely to access a diverse range of sources when retrieving news through digital intermediaries, such as search engines and social media (Fletcher et al., 2023). Digital intermediaries also tend to provide similar news, and results are not heavily personalised (Nechushtai and Lewis, 2019; Nechushtai et al., 2023). Overall, people are likely to see a rich variety of news sources through online intermediaries, and results are likely to be similar for most people.

This report furthers research focused on understanding how search engines affect the news content encountered by ordinary internet users. Our summary of findings draws on data obtained through the ADM+S Australian Search Experience ‘citizen science’ research project, to understand what news is presented to Australians through Google News. Data collection concluded in late July 2022. In contrast to other studies, we draw on a wider set of search terms (Nechushtai and Lewis, 2019; Nechushtai et al., 2023) and offer a more comprehensive account of Google News (Fletcher et al., 2023).

The purpose of this research project was to investigate the recommendations made by leading search engines for a wide range of search topics. We collected ‘data donations’ from a self-selected sample of the Australian general public via a browser plug-in available for the desktop versions of web browsers, including Google Chrome, Mozilla Firefox and Microsoft Edge.²

² The plug-in simulates web browsing sessions for mobile devices as part of data collection. We have incorporated mobile data into our results.



3. METHOD

When installed, the plug-in runs a process at four-hour intervals, provided that the computer is running and the web browser is open. Our plug-in borrows the user’s online persona and then initiates a series of queries to the Google News search engine, among other platforms.³ The first page returned by the query, yielding an arbitrary number of individual search results, constitutes the ‘data donation’, which is then submitted to a central online database.

The plug-in cycles through a set of search terms that can be varied over the project’s lifetime. These searches will appear as if the user had initiated them themselves, thereby producing organic search results. Our initial search queries included the names of key politicians and political parties, as well as terms related to current Australian and world events (e.g., COVID-19 and climate change).

The search results were restricted to the period 1 September 2021 – 1 January 2022, Australia Eastern Standard Time (Brisbane). During this period, 401 participants submitted data donations for the sample dataset. While the results are not representative of the entire Australian population (see Appendix), they are statistically significant.⁴

Search terms were selected to inform the analysis. Several of these terms were aligned with politics, which allowed us to establish the ‘Australian Politics’ keyword set; other keywords were aligned with international issues, which were grouped under the ‘Global Issues’ keyword set (see Table 1).

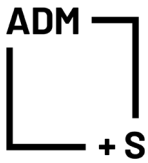
Australian Politics Search Terms	Global Issues Search Terms
Adam Bandt, Anthony Albanese, Barnaby Joyce, Greens, Labor Party, Liberal Party, National Party, One Nation, Pauline Hanson, Scott Morrison	Afghanistan, COP26, COVID, Feminism, Lockdown, Quarantine, Renewable energy, Tokyo 2021, Tokyo Olympics, Travel rules, Vaccine

Table 1: Search queries used for the study.

It is important to note that the research was query driven and that the search queries varied significantly. Accordingly, we cannot claim to offer a complete account of what Google News presents. However, the ‘Australian Politics’ keyword set captures major actors in the political

³ These are search queries rather than topic queries, which are also possible on Google News.

⁴ The results rejected the null hypothesis, and statistical significance was achieved for the central tendencies of all distributed nominal data within the confidence interval of the 95th percentile.



sphere at the time of data collection, and therefore, it can be viewed as a robust attempt to capture topical news stories. Conversely, though we attempted to select internationally relevant topics, there are several newsworthy issues occurring at any one time, and therefore, the results for the 'Global Issues' keyword set should be viewed with caution.

4. RESEARCH APPROACH

For our analysis, we focused on two key aspects of exposure diversity. First, we attempted to identify the dominant news sources in our dataset. To do this, we observed how many times a domain (or website) was returned in the search results for each keyword.

We did not specifically focus on the ranking of the results for this analysis, so a positive count could mean that the website was possibly the top-ranked search result or, alternatively, the last on the page (10th). We found that the position of news sources shifted erratically, and we could not establish visible patterns in these rankings. Accordingly, we are unable to report meaningful trends with respect to this statistic.

Second, we focused on geography and assessed how much Australian news was being presented to Australians. For this second assessment, we established two hypotheses.

Hypothesis 1: 'Australian politics' search terms will be localised, and Australian news sources will be prominent.

Hypothesis 2: Search terms related to global politics will return international news sources.

To answer these hypotheses, we identified countries of operation by capturing the country in which the domain was registered. Where this was not possible, we used the country-specific domain suffix (e.g., 'www.news.com.au' has the country suffix 'au').

We also had a significant number of results and a noticeable long tail, so we only focused on the top 300 returning domains to ensure the results were legible and relevant.

Consistent with previous work on the Australian Search Experience, we also assessed whether the Google News results were personalised in any way. To do this, we explored whether different user demographics were directed to different news sources (e.g., based on assumed geographic location or political leaning), whether any patterns varied with the theme of the search and whether the mix of sources changed over time.



5. RESULTS

Personalisation

Our preliminary results show **limited evidence of personalisation** in Google News. This aligns with existing findings in the international literature, as well as with previous work conducted by ADM+S concerning Google Search (Haim et al., 2018; Nechushtai and Lewis, 2019; Nechushtai et al., 2023; Bruns, 2022).

For example, Figure 1 shows that, for the keyword ‘COVID’, there are few substantial differences between the sources recommended to participants with different party preferences; Labor supporters are somewhat more likely to encounter *New York Times* content and less likely to encounter British Broadcasting Corporation (BBC) content, whereas non-aligned participants encounter more content from *Guardian Australia* compared with content from the ABC. Coalition supporters’ recommendations diverge somewhat more strongly from overall patterns, but rather than receiving disproportionately more recommendations from conservative news sources, they are largely directed to *Guardian Australia* and BBC content.

News Domains by Party Preference

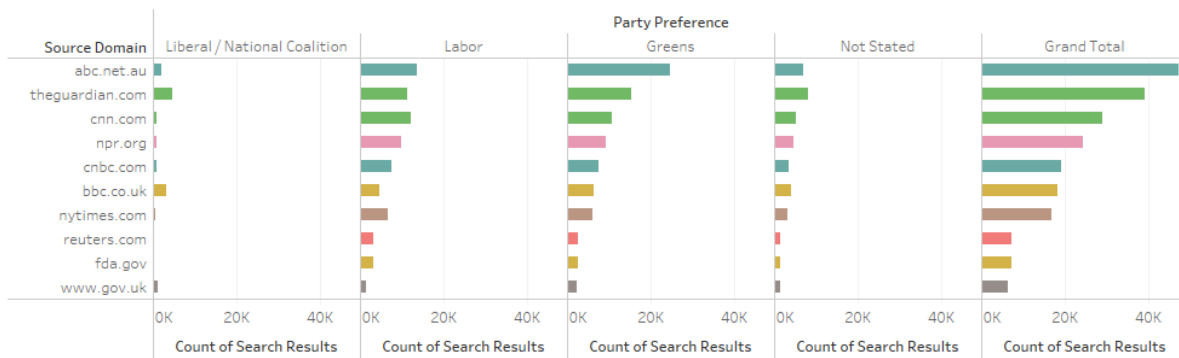


Figure 1: Most common news sources for the keyword ‘COVID’, by party preference.

Given that Google News results are not subject to deep personalisation, we can now consider the frequency of results and location, as these results will hold for the entire cohort.



Frequency

An isolated subset of domains frequently occupied the first page of search results for **Google News**. Certain news sources were prominent across the entire data collection period, and Google News indexed them quite well.

For both keyword sets ('Australian Politics' and 'Global Issues'), *Guardian Australia* was consistently the most frequently observed news website, followed by the ABC.

To better understand our results, we split them into quartiles and found that, for both keyword sets, seven domains took up ~25% of the total of their respective sample.⁵ As shown in Figure 2, several major Australian news organisations were regularly observed in the 'Australian Politics' keyword set results (represented by the blue quadrant).

The tables below shows how regularly these domains were observed as a percentage of the top 300 domains.

This initial analysis also revealed differences between each keyword set. The 'Global Issues' keyword set had a significantly higher proportion of international news outlets (e.g., The Guardian (U.K.), CNN, *The New York Times* and *Reuters*), whereas the 'Australian Politics' keyword set had a variety of Australian news outlets, ranging from public service media outlets such as the ABC to the News Corp–owned Sky News Australia.

'Australian Politics'	'Global Issues'
1. theguardian.com (6.85% of 300 domains)	1. theguardian.com (5.02% of 300 domains)
2. abc.net.au (4.03%)	2. abc.net.au (4.74%)
3. smh.com.au (3.98%)	3. renew.biz (4.52%)
4. dailymail.co.uk (3.93%)	4. nytimes.com (3.27%)
5. news.com.au (3.25%)	5. cnn.com (2.69%)
6. skynews.com.au (2.53%)	6. environmentalleader.com (2.55%)
7. nzherald.co.nz (2.44%)	7. reuters.com (2.53%)

Table 2: Proportion of news outlets observed for each keyword.

By further breaking down the percentages *within* the top seven domains, we were able to examine the observation frequency in this quartile.

⁵ This was not an exact quartile but the closest we could obtain (e.g., 24%–28%), given the variability of the underlying data.



'Australian Politics'	'Global Issues'
1. theguardian.com (AU) (25.35% of the top seven domains)	1. theguardian.com (U.K.) (19.83% of the top seven domains)
2. abc.net.au (14.94%)	2. abc.net.au (18.72%)
3. smh.com.au (14.75%)	3. renew.biz (17.84%)
4. dailymail.co.uk (14.54%)	4. nytimes.com (12.9%)
5. news.com.au (12.02%)	5. cnn.com (10.62%)
6. skynews.com.au (9.35%)	6. environmentalleader.com (10.09%)
7. nzherald.co.nz (9.05%)	7. reuters.com (9.99%)

Table 3: Observation frequency within top seven domains.



Figure 2: Frequency of observations across the 'Australian Politics' keyword set.



There is less diversity when we focus on specific keywords. For example, the top quartile of results for the keyword 'Barnaby Joyce' is split between *Guardian Australia* and *news.com.au* (see Figure 3). The keyword 'COVID' predominantly returned news from *abc.net.au*, followed by *The Guardian (U.K.)* and CNN (see Figure 4).

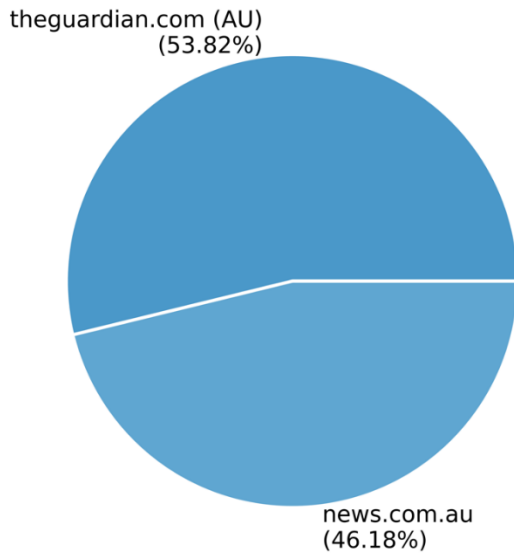


Figure 4: Top quartile (26.77%) of results for the keyword 'Barnaby Joyce'.

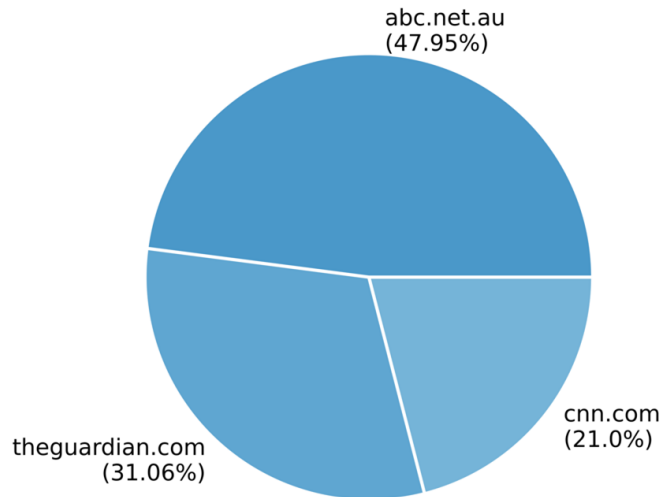


Figure 4: Top quartile (27.86%) of results for the keyword 'COVID'.

Location

We expected that the 'Australian Politics' keyword set would localise and provide Australian news and that the 'Global Issues' keyword set would return more international news.

Our Australian Politics search terms did not always localise. For example, if one searches for 'National Party' in Australia, one would expect news about the National Party of Australia. However, we found many domains from New Zealand about its National Party. This provides further evidence of the lack of personalisation in Google News and the growing importance of search queries when finding content. **Names were more reliable**, and search queries for politicians (e.g., 'Scott Morrison') resulted in a greater representation of Australian content.

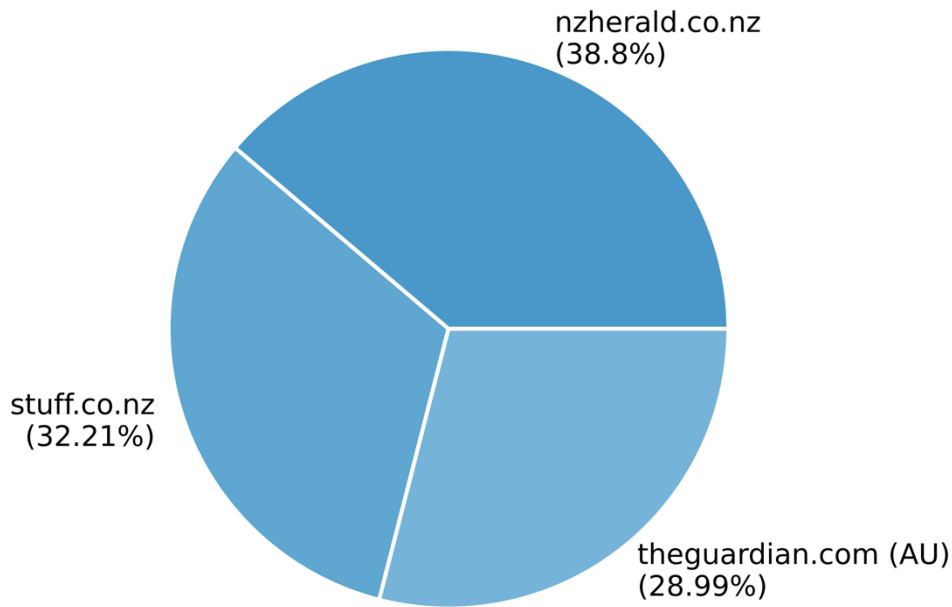


Figure 5: Top quartile (28.1%) of results for the keyword 'National Party'.

Despite these differences, the most frequently observed domains for the 'Australian Politics' keyword set overall were predominantly Australian.

'Australian Politics'
1. theguardian.com (Australia)
2. abc.net.au (Australia)
3. smh.com.au (Australia)
4. dailymail.co.uk (United Kingdom/Australia)
5. news.com.au (Australia)
6. skynews.com.au (Australia)
7. nzherald.co.nz (New Zealand)

Table 4: Frequently observed domains for the 'Australian Politics' keyword set.

The results from the 'Global Issues' keyword set corroborated our original hypothesis: The only Australian source in the top quartile was the Australian Broadcasting Corporation, and the rest were international sources. There was evidence of a preference for domains from the United States and the United Kingdom.

The top results for the keyword 'Travel rules' illustrate this growing international presence (see Figure 6).

Some of the keywords focused on sustainability (e.g., 'renewable energy' and 'COP26'), which explains the presence of specialised environmental sources in the top results.



It remains an open question whether queries for global news topics, such as ‘vaccines’, ‘lockdown’ or ‘climate change’, should preference local sources.

‘Global Issues’	
8.	theguardian.com (United Kingdom)
9.	abc.net.au (Australia)
10.	renew.biz (N/A)
11.	nytimes.com (United States)
12.	cnn.com (United States)
13.	environmentalleader.com (N/A)
14.	reuters.com (United Kingdom/International)

Table 5: Results from ‘Global Issues’ keyword set

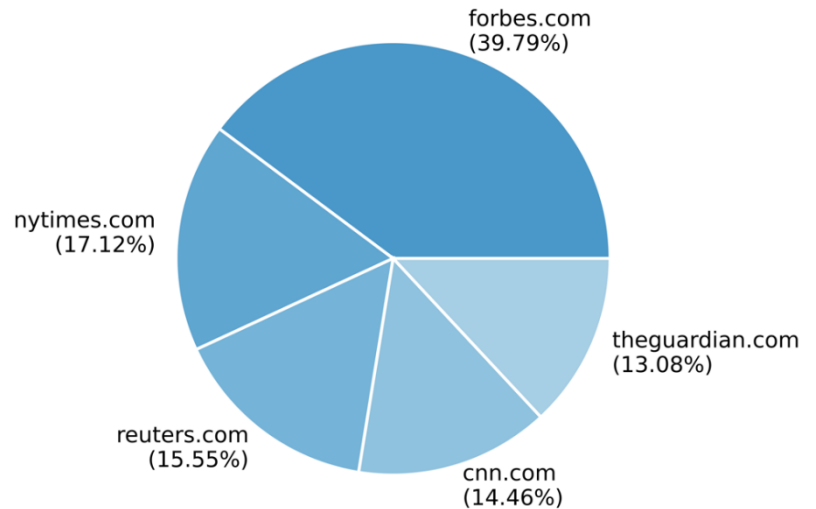
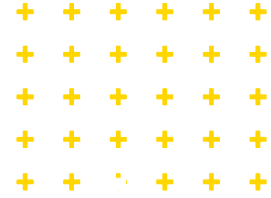


Figure 6: Top quartile (27.73%) of results for the keyword ‘Travel rules’.

Limitations

These results relate to highly specific search terms queried within a time-bound period. People have significantly diverse search patterns, and different queries naturally produce different results.

We cannot account for external factors that may inform the frequency count. For example, if a news outlet produces more content about a particular politician or issue, it is possible Google will index more stories.



6. APPENDIX: ADDITIONAL DETAILS ON METHOD

Any single search result conveys a variety of attributes that include but are not limited to the search topic, the discrete numerical ranking within the list of alternative search results (i.e., the data donation), the retrieval time, the URL of respective news articles, the publisher name of the news outlet and the unique identifier relating the search result to its contributing participant.

The 401 participants had a variety of characteristics. Language had also been studied but was disregarded, since the majority of participants reported 'English' as their spoken language.

Figure 7 illustrates the demographics in terms of sample numbers

7. REFERENCES

Bruns, A. (2022) Background paper: Australian search experience project, ADM+S working paper series 2022 (001), ARC Centre of Excellence for Automated Decision-Making and Society, <https://doi.org/10.25916/k7py-t320>

Fletcher, R., Kalogeropoulos, A. & Nielsen, R. K. (2023). More diverse, more politically varied: How social media, search engines and aggregators shape news repertoires in the United Kingdom. *New Media & Society*, 25(8), 2118–2139.

Haim, M., Graefe, A., & Brosius, H. B. (2018). Burst of the filter bubble? Effects of personalization on the diversity of Google News. *Digital Journalism*, 6(3), 330–343.

Helberger, N. (2019). On the democratic role of news recommenders. *Digital Journalism*, 7(8), 993–1012.

Napoli, P. M. (2011). Exposure diversity reconsidered. *Journal of Information Policy*, 1, 246–259.

Nechushtai, E., & Lewis, S. C. (2019). What kind of news gatekeepers do we want machines to be? Filter bubbles, fragmentation, and the normative dimensions of algorithmic recommendations. *Computers in Human Behavior*, 90, 298–307.

Nechushtai, E., Zamith, R., & Lewis, S. C. (2023). More of the same? Homogenization in news recommendations when users search on Google, YouTube, Facebook, and Twitter. *Communication and Society*, 1–27.

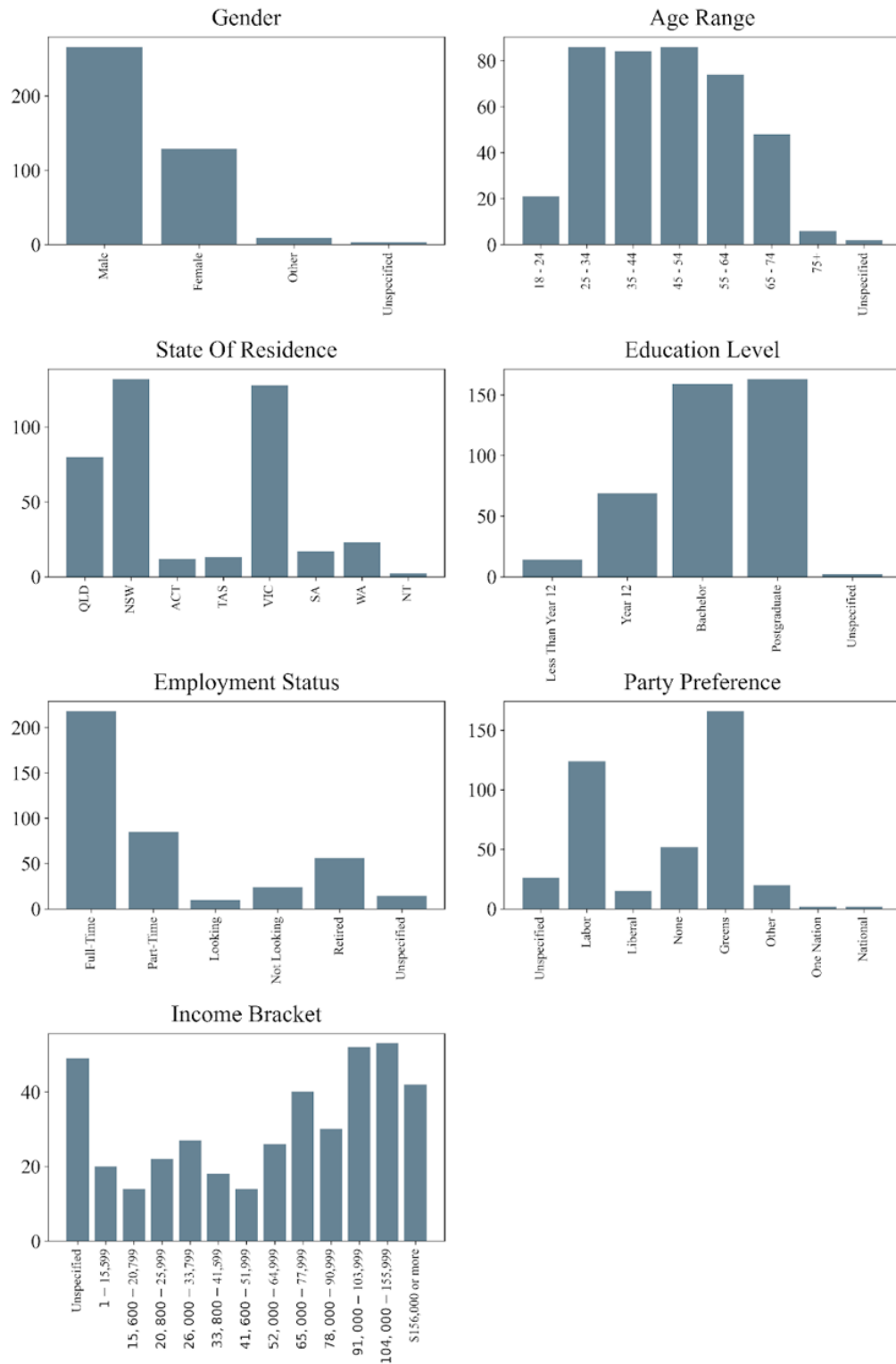


Figure 7: Participants' demographic data.



admscentre.org.au



Australian Government
Australian Research Council

This Centre is funded by the
Australian Government
through the Australian
Research Council

