Australian Code of Practice on Disinformation and Misinformation

Google Annual Transparency Report, May 2022

1st January 2021 - 31st December 2021

Introduction

As Google's mission is to organise the world's information to make it universally accessible and useful, combating misinformation and disinformation is of utmost importance. This is especially so around issues such as public health, elections, civic engagement, or other issues that directly impact our users' lives and civil society.

Since Google's founding, our product, policy, and content enforcement decisions have been guided by the following three principles:

1. **Value openness and accessibility:** Aim to provide access to an open and diverse information ecosystem, while maintaining a responsible approach to supporting information quality;

2. **Respect user choice:** Based on users' intent, access to content that is not illegal or prohibited by our policies should be available, but set a high bar for information quality where users have not clearly expressed what they are looking for;

3. **Build for everyone:** Take into account the diversity of users (cultures, languages, backgrounds) and seek to address their needs appropriately.

With these principles in mind, we implement a multi-faceted approach to address the complex challenges and risks raised by misinformation and disinformation across many of our products and services.

This report discusses these approaches across many of Google's diverse products and services, particularly in the context of addressing misinformation and disinformation. The contents of the report are organised through the framework of Objectives 1 through 7 of the Code, and cover the period from 1 January 2021 to 31 December 2021. We highlight relevant policies, product features and tools, and enforcement efforts across Google's key consumer-facing information services, such as Google Search, Google News, Google Ads, Google AdSense and YouTube. Unless stated otherwise, all policies apply to users and content creators in Australia.

We expect to publish subsequent versions of this report annually. We are also committed to improving and augmenting future iterations with further insights relevant to our continued efforts to combat misinformation and disinformation.

With respect to terminology, we acknowledge the differences between 'misinformation' and 'disinformation', as outlined by this Code. The main difference pertains to whether the individual creating or sharing content has an *intent to deceive*. Since we cannot definitively identify this intent and distinguish between misinformation and disinformation, we use the term 'misinformation' for brevity and focus instead on the specific behaviours or types of content that we seek to either prohibit, discourage or reward through our policies and enforcement actions.

Commitments under the Code

Google, including YouTube, has committed to all 7 Objectives and related Outcomes provided in the Code and detailed below. As we respond and adapt to new and evolving challenges relating to misinformation, we continually review our products, policies, enforcement and transparency work to ensure that they are as effective as possible.

Objective and Outcomes
Objective 1 - Safeguards against Disinformation and Misinformation
1a   Signatories contribute to reducing the risk of Harms that may arise from the propagation of Disinformation and Misinformation on digital platforms by adopting a range of scalable measures.
1b   Users will be informed about the types of behaviours and types of content that will be prohibited and/or managed by Signatories under this Code.
1c   Users can report content or behaviours to Signatories that violate their policies under section 5.10 through publicly available and accessible reporting tools.
1d   Users will be able to access general information about Signatories' actions in response to reports made under 5.11.
Objective 2 - Disrupt advertising and monetisation incentives for disinformation
2   Advertising and/or monetisation incentives for Disinformation are reduced.
Objective 3 - Work to ensure the integrity and security of services and products delivered by digital platforms.
3   The risk that Inauthentic User Behaviours undermine the integrity and security of services and products is reduced.
Objective 4 - Empower consumers to make better informed choices of digital content.
4   Users are enabled to make more informed choices about the source of news and factual content accessed via digital platforms and are better equipped to identify Misinformation.
Objective 5 - Improve public awareness of the source of political advertising carried on digital platforms.
5   Users are better informed about the source of Political Advertising. Signatories detail policies that provide users with information about the source of Political Advertising carried on digital platforms.
Objective 6 - Strengthen public understanding of Disinformation and Misinformation through support of strategic research.
6   Signatories support the efforts of independent researchers to improve public understanding of Disinformation and Misinformation.
Objective 7 - Signatories will publicise the measures they take to combat Disinformation.
7   The public can access information about the measures Signatories have taken to combat Disinformation and Misinformation.

Objective 1: Safeguards against Disinformation and Misinformation

Outcome 1a: Signatories contribute to reducing the risk of Harms that may arise from the propagation of Disinformation and Misinformation on digital platforms by adopting a range of scalable measures.

Across many of our products and services, we deploy a range of measures to address the risk of potential harm caused by the propagation of misinformation. We present a number of outcome-focused metrics to this effect below.

Actions taken to address coordinated influence operations

Our Threat Analysis Group and Trust & Safety teams work to monitor malicious actors around the globe, disable their accounts, and remove the content that they posted – including but not limited to operations that may affect Australia. We provide monthly updates about our actions against coordinated influence operation campaigns in our Threat Analysis Group blog (see below for examples of actions from 2021).

Case Study: Examples of actions against coordinated influence operations from 2021

- In March 2021, we **terminated 33 YouTube channels** and **2 advertising accounts** as part of our ongoing investigation into coordinated influence operations linked to Myanmar. This campaign uploaded content in Burmese about the military coup in Myanmar.

- In June 2021, we **terminated 17 YouTube channels** and **blocked 1 domain from eligibility to appear on Google News surfaces and Discover** as part of our investigation into coordinated influence operations linked to Ukraine. This campaign uploaded content in Ukrainian that amplified several media platforms posing as news outlets and promoting a select number of local politicians.

- In October 2021, we **terminated 4 AdSense accounts** and **blocked 22 domains from eligibility to appear on Google News surfaces and Discover** as part of our investigation into a reported coordinated influence operation linked to India. The campaign uploaded a variety of news content in English to domains that were designed to look as if they were independent news outlets in various US states and European countries.

Enforcement of YouTube's Community Guidelines

As detailed in our [Community Guidelines](), YouTube does not allow misleading or deceptive content that poses a serious risk of egregious harm. We enforce our policies across the globe, including in Australia, using a combination of content reviewers and machine learning to remove content that violates our policies as quickly as possible. More information on how YouTube addresses misinformation can be found [here]().

In 2021, the following actions were taken:

25,000,000+ — YouTube **videos removed globally** for violating the Community Guidelines

90,000+ — YouTube **videos** that violated Community Guidelines and were uploaded from IP addresses in **Australia** were **removed**.

700,000+ — YouTube **videos** were **removed globally** as content related to dangerous or misleading COVID-19 information.

5,000+ — YouTube **videos** uploaded from IP addresses in **Australia** and where content related to dangerous or misleading COVID-19 information were **removed**.

As mentioned above, we rely on a combination of people and technology to flag inappropriate content and enforce these guidelines. These flags can come from our automated flagging systems, from members of the [Trusted Flagger program]() or from users in the broader YouTube community.

2021 Violative View Rate Estimate

To measure our progress on removing violative videos, we have developed a metric called Violative View Rate (VVR). This metric is an estimate of the proportion of video views that violate our Community Guidelines in a given quarter (excluding [spam]()). In order to calculate VVR, we take a sample of the views on YouTube and send the sampled videos for review. Once we receive the decisions from reviewers about which videos in the sample are violative, we aggregate these decisions in order to arrive at our estimate. Additional data on VVR can be found [here](). A more detailed explanation of the VVR calculation (including which policies are included) is found within the [Community Guidelines Report FAQs](), and further information on the VVR methodology can be found [here]().

- In **Q1** (Jan - Mar 2021), VVR was **0.16 - 0.18%** (i.e., out of every 10,000 views on YouTube, 16-18 were of violative content).
- In **Q2** (Apr - Jun 2021), VVR was **0.19 - 0.21%** (i.e., out of every 10,000 views on YouTube, 19-21 were of violative content).
- In **Q3** (Jul - Sep 2021), VVR was **0.09 - 0.11%** (i.e., out of every 10,000 views on YouTube, 9-11 were of violative content).

- In **Q4** (Oct - Dec 2021), VVR was **0.12 - 0.14%** (i.e., out of every 10,000 views on YouTube, 12-14 were of violative content).

More information on metrics pertaining to content that violates our Community Guidelines can be found in the [YouTube Community Guidelines Enforcement Report](#).

Outcome 1b: Users will be informed about the types of behaviours and types of content that will be prohibited and/or managed by Signatories under this Code.

We aim to ensure that our policies across products and services are available to the public, users and creators in a form that is clear, predictable and repeatable. Each of our product policies address the types of behaviours and content prohibited on the product, with examples as needed. A list of product-specific policies are available on [this page](#), while those most relevant for this report are listed below:

Google Search & Google News
- [Google Search Webmaster Guidelines](#)
- The [Google Search Content policies](#) include policies related to Search Features, such as:
    - Deceptive practices policy
    - Manipulated media policy
    - Medical content policy
- Content on Google News must follow all Google Search Content policies, as well as [Google News-specific policies](#) which include:
    - Ads & sponsored content policy
    - Misleading content policy
    - Transparency policy

Google Ads
- [Google Ads Policies](#) cover four broad areas:
    - Prohibited content
    - Prohibited practices
    - Restricted content and features
    - Editorial and technical requirements
- [Google Ads Misrepresentation policies](#) include, but are not limited to:
    - Unacceptable business practices policy
    - Coordinated deceptive practices policy
    - Misleading representation policy
    - Manipulated media policy
    - Unreliable claims policy
- [Google Ads Inappropriate content policy](#)

Google AdSense

- All AdSense publishers are required to follow [Google Publisher policies](#), which include rules against misrepresentative content.
- [Google AdSense Program policies](#) include policies regarding:
    - Encouraging clicks or views
    - Ad behaviour
    - Deceptive site navigation.

YouTube

- [YouTube Community Guidelines](#)
    - [YouTube misinformation policies landing page](#), includes links to:
        - [Misinformation policies](#)
        - [Election misinformation policies](#)
        - [COVID-19 medical misinformation policy](#)
        - [Vaccine misinformation policy](#)
    - [YouTube Impersonation policy](#)
    - [YouTube fake engagement policy](#)
    - [YouTube spam, deceptive practices, and scams policies](#)
- [YouTube advertiser-friendly content guidelines](#)
- [YouTube channel monetisation policy](#)

Further details related to these policies can be found in *[Appendix A](#)*.

Outcome 1c: Users can report content or behaviours to Signatories that violates their policies under section 5.10 through publicly available and accessible reporting tools.

Our products and services provide publicly available and accessible channels or mechanisms that allow users to report content that they believe has violated our policies. Highlighted below are examples of reporting mechanisms that users can utilise to provide feedback across Google Search, Google Ads, Google AdSense and YouTube.

In **Google Search**, users can provide feedback on an overall Search results page or on specific features such as Autocomplete predictions or Knowledge Panels.

Figure 1c.i. '*Send Feedback' button in a Knowledge Panel*
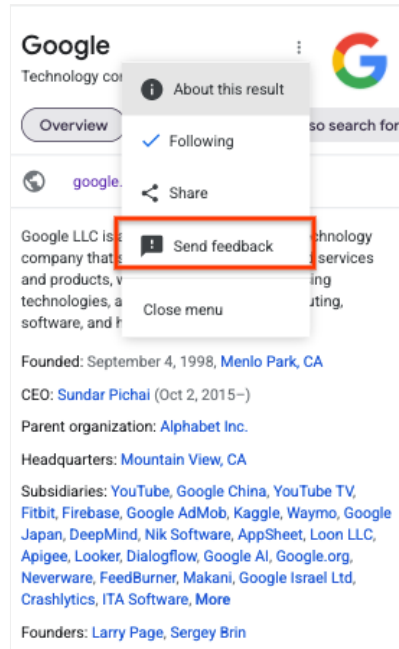
In **Google Ads,** users can tap the 'Why this Ad' button on the top right corner of an ad and select 'Report this Ad' to let us know about ads that they believe violate our policies.

Figure 1c.ii. Opening the 'why this ad' interface and 'Report this ad' button
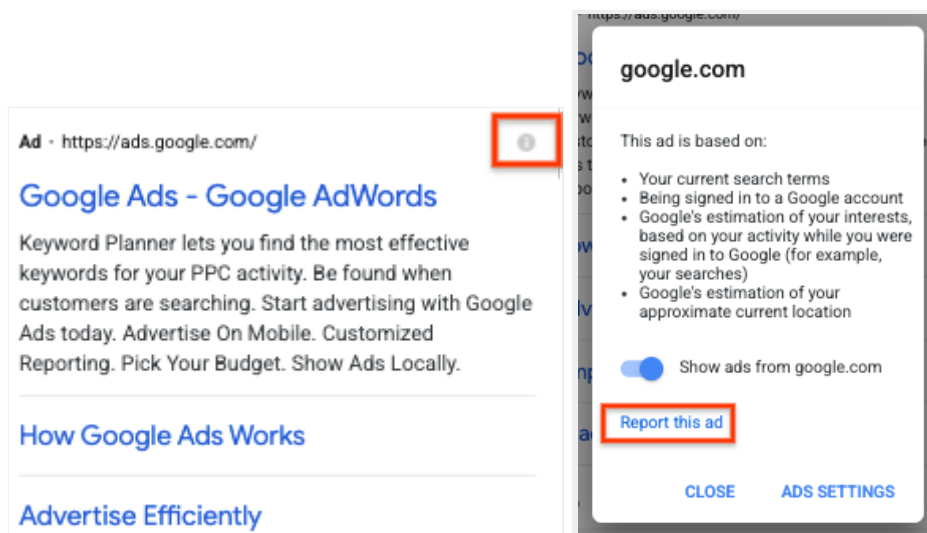
Clicking on 'Report this ad' will redirect users to the 'Report an ad/listing' form. Using this feature, individuals can choose the reason that best describes their complaint.

Figure 1c.iii. *'Report an ad/listing' form for users*

On **Google AdSense**, users can report a site that they believe is showing ads in violation of our product policies.

Figure 1c.iv. *'Reporting a Violation' form on Google AdSense*

On **YouTube**, users can utilise our 'Report' button to flag issues with any video and indicate information about which of our policies they believe the video is violating. Users can also report inappropriate channels, playlists, comments and other content. Further details are available on this page.

Figure 1c.v. *'Report' button on YouTube user interface*

Outcome 1d: Users will be able to access general information about Signatories' actions in response to reports made under 5.11.

We provide regular, publicly available reports on enforcement of our content policies - these include information regarding actions in response to reports of misinformation. A list of these reports (including public links to the materials) is included below. Information from these sources have also been incorporated throughout the relevant sections of this report.

- Our [Google Transparency Report website](#) is a centralised hub for transparency reporting on key content topics across various Google products and services;

- The annual [Ads Safety Report](#) provides updates on policy enforcement in Google Ads;

- Our Threat Analysis Group Quarterly Bulletin (published on our [Threat Analysis Group blog](#)) discloses actions we have taken against coordinated influence operation campaigns on our platforms;

- Our [YouTube Community Guidelines Enforcement report](#) provides a quarterly update on the work we do to enforce our policies on YouTube.

Objective 2: Disrupt advertising and monetisation incentives for disinformation

Outcome 2: Advertising and/or monetisation incentives for Disinformation are reduced.

Relevant policies across our products and services

Our advertising and monetisation policies prohibit a range of behaviours and types of content that are clearly connected to misinformation, or that commonly overlap with misinformation. These policies include, but are not limited to:

- Google Ads Policies
- Google Publisher Policies (any user monetising content with Google ad code)
- YouTube advertiser-friendly content guidelines
- YouTube channel monetisation policy

Further details on these policies can be found in *Appendix A*.

Updating our monetisation policies related to misinformation

As the misinformation landscape and the narratives that propagate through it constantly evolve, our efforts and interventions to combat misinformation must adapt accordingly. This includes updating our policies and monitoring risks associated with misinformation in the context of broad societal issues that impact our users' lives, such as climate change. We have updated our Ads, AdSense and YouTube policies, as explained below.

Updating our monetisation policies on climate change

Due to the growing number of advertising and publisher partners that have expressed concerns about ads promoting inaccurate claims concerning climate change, we announced a new addition to our monetisation policies for Google advertisers, publishers and YouTube in October 2021. This policy addition prohibits ads for, and monetisation of, content that runs contrary to scientific consensus around the existence and causes of climate change. More specifically, this includes prohibiting:

- Content referring to climate change as a hoax or scam,
- Claims denying that long-term trends show the global climate is warming, and
- Claims denying that greenhouse gas emissions or human activity contribute to climate change.

Further information regarding this update and links to these policies can be found here and In our Policy Help Centers.

Google Ads also provides its advertising partners with features that enable them to maintain control over where their ads appear, the format in which their ads run, and their intended audience. In addition to updating our monetisation policies in April 2021, Google Ads developed a new feature called Dynamic Exclusion Lists, which helps our advertisers to seamlessly and continuously prevent ads from serving alongside certain content. Further information can be found here.

Enforcing our policies to reduce monetisation incentives for misinformation
To ensure that advertisers and publishers on our networks are complying with our policies, we continuously monitor our advertising networks and use a combination of algorithmic and human reviews.

Our 2021 Ads Safety Report and Ads Safety blog post for 2021 are publicly available and contain data that exemplify the enforcement actions taken on ads and publisher content. We have included key global metrics from these sources, as well as Australia-specific enforcement data for violation of misrepresentation ads policies, below.

3.4 billion	**'bad ads'** blocked or removed for policy violations
652.1 million	**ads** blocked or removed for 'Abusing the Ad Network'
38.1 million	**ads** blocked or removed for violating our misrepresentation policies
1.7 billion	**pages** in our publisher network with ads blocked or restricted for violating our policies
63,000	**publisher sites** where ads were blocked or restricted for violating our policies
500,000+	**pages** with blocked ads that violated our policies against harmful health claims related to COVID-19 and demonstrably false claims that could undermine trust and participation in elections
5.6 million	**advertiser accounts** suspended for policy violations
657,000+	**creatives** blocked from Australia-based advertisers for violating our misrepresentation ads policies (misleading, clickbait, unacceptable business practices, etc.)

Objective 3: Work to ensure the integrity and security of services and products delivered by digital platforms.

Outcome 3: The risk that Inauthentic User Behaviours undermine the integrity and security of services and products is reduced.

Google is continually working to address and mitigate risks associated with behaviours that seek to undermine the integrity and security of our products and services ('Inauthentic User Behaviours'). Detailed below are the targeted policies that we have put in place across our products and services, tailored to the specific risks faced by each product or service in relation to Inauthentic User Behaviour.

Google Search & Google News

Our Google Search webmaster guidelines prohibit the following techniques, which may be misused to deceive our ranking systems or users:

- Automatically generated content
- Participation in link schemes
- Cloaking
- Sneaky redirects
- Hidden text or links
- Doorway pages
- Scraped content
- Loading pages with irrelevant keywords
- Creating pages with malicious behaviour
- Abusing structured data markup
- Sending automated queries to Google

Where a website or page result violates one or more of these guidelines, Google may act against it. Sites may be submitted for reconsideration once the issue is remedied.

The Google Search Content policies include policies related to Search Features, such as:

- Deceptive practices policy
- Manipulated media policy
- Medical content policy

Content on Google News must follow all Google Search Content policies, as well as Google News-specific policies which include:

- Ads & sponsored content policy

- Misleading content policy
- Transparency policy

Google Ads

Ads that deceive users by excluding relevant product information or providing misleading information about products, services, or businesses are prohibited on Google. To minimise Inauthentic User Behaviour, the Ads misrepresentation policy highlights unacceptable practices and behaviours that advertisers must avoid.

Google AdSense

Google enables a free and open web by supporting publishers monetising their content and advertisers reaching prospective customers with useful, relevant products and services. If users wish to monetise their content with Google ad code, they are required to adhere to the AdSense Program policies which prohibit:
- Invalid clicks and impressions;
- Encouraging clicks or views (non-rewarded inventory);
- Ad placement on pages receiving traffic from certain sources;
- Modification of AdSense code that affects ad behaviour; and
- Inappropriate ad placement.

Case Study: Invalid Traffic

Google has developed processes and invested in technology to remove invalid traffic and advertising fraud from our systems. Invalid traffic, as defined in Google AdSense policies, refers to clicks and impressions on ads that are not a result of genuine user interest. This could include manual clicks meant to increase someone's advertising costs or meant to increase profits for website owners hosting ads. Additionally, invalid traffic also includes clicks and impressions by automated tools, robots, or other deceptive software.

Our Ad Traffic Quality Team uses live reviewers, automatic filters, machine learning, and research to detect and filter as much invalid and fraudulent activity as possible. We monitor and analyse traffic to prevent advertisers from paying for invalid clicks, impressions, views, or interactions, and stop publishers that generate invalid activity from receiving undeserved advertising income. Our efforts protect advertiser spend, enabling advertisers to maximise the return on their investment, while also ensuring that legitimate publishers have a level playing field that increases their chances to monetise their content. Additional information on how Google protects users from invalid activity can be found here.

YouTube

Our YouTube Community Guidelines include policies to prohibit content intended to impersonate a person or channel. These policies apply to all types of content on the platform, including videos, comments and links. Below we have listed relevant policies:
- Impersonation policy
- Fake engagement policy
- Spam, deceptive practices, and scams policies

More information related to 'Inauthentic User Behaviours' can be found in the Threat Analysis Group case study (see Outcome 1a) and advertisement policy enforcement metrics (see Outcome 2).

Objective 4: Empower consumers to make better informed choices of digital content.

Outcome 4: Users are enabled to make more informed choices about the source of news and factual content accessed via digital platforms and are better equipped to identify Misinformation.

Tools and features enabling users to make more informed choices

Across many of our products, we provide users with a variety of opportunities to make informed choices about content encountered, thereby allowing users to identify misformation. With more tools to identify misinformation, users are then empowered to report instances of misinformation using the mechanisms outlined in Outcome 1c, thus helping reduce the overall risk of harm from misinformation. Some examples of tools and features created to elevate authoritative sources and to help users make informed choices about content are outlined below:

- Fact-check snippets and tags on Google Search, Images, and News;
- Elevating original reporting in Google Search;
- 'Full Coverage' feature in Google News;
- 'Information Panels', including Fact Check, and those giving publisher context and topical context on YouTube (further information here);
- 'Breaking News' and 'Top News Shelves' from authoritative sources on YouTube;
- Raising authoritative voices for newsworthy events and topics prone to misinformation in YouTube search results and recommended videos (further information here);
- Labelling state-funded news channels on YouTube.

A detailed explanation of these tools and features is found in *Appendix B*. Note that these tools and features are automatically available to all users globally, including those in Australia, and their availability does not require an individual user to select and/or activate them.

Ranking Google Search results

We continue to improve the design of our ranking systems, which are key to helping users make informed decisions and reducing the proliferation of misinformation. For example, Google Search's algorithms consider a host of 'signals' (or characteristics of a web page) that are indicative of high-quality and reliable information. To ensure Google Search algorithms meet high standards of relevance and quality, we have a rigorous process that involves both live tests and thousands of trained external Search Quality Raters from around the world. Raters do not determine the ranking of

an individual, specific page or website, but they help us benchmark the quality of our results so that we can meet a high bar for users of Google Search all around the world. Under our [Search Quality Rater Guidelines](), raters are instructed to assign the lowest rating to pages that are potentially harmful to users or specified groups, misleading, untrustworthy, and spammy. Examples of such pages include those that contain clearly inaccurate harmful information that can easily be refuted by straightforward and widely accepted facts; harmful information that contradicts well-established expert consensus; and harmful, unsubstantiated theories/claims not grounded in any reasonable facts or evidence.

In addition, these guidelines include very high Page Quality rating standards for what we call 'Your Money or Your Life' pages, which describe topics that could potentially impact a person's future happiness, health, financial stability or safety. These include areas like: civics, government, law, finance, shopping, health and safety, etc.

Reliable information on Google Search

Google Search systems have been trained to detect when a topic is rapidly evolving and a range of sources have not yet weighed in. In such cases, users will be prompted with a notice indicating that it may be best to check back later when more information from a wider range of sources might be available, as shown in *Figure 4.i.* below (more information found [here]()).

*Figure 4.i.* Notice to users about a rapidly evolving topic
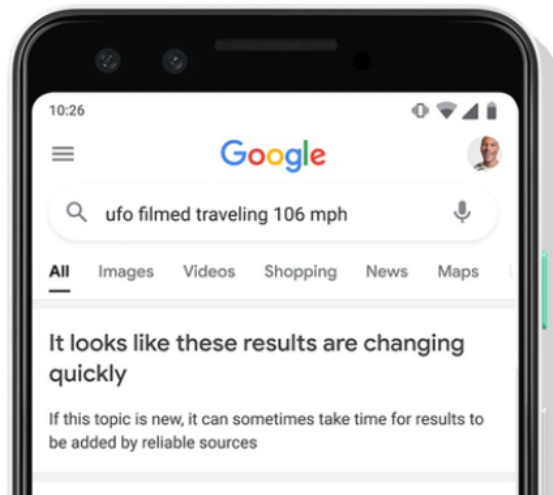
We also recently launched the 'About This Result' feature, which enables users to quickly find information about sources on Google Search, as shown in *Figure 4.ii* below (more information found [here]()). Additional context about Google Search results helps users:

- Learn about the source to get a sense of how they describe themselves, in their own words;
- Learn what others on the web have said about a site;
- Learn additional context about the topic, like top news coverage, to see what a range of sources have to say.

With this additional context, users can make a more informed decision about the sites they may want to visit and what results will be most useful for them.

Figure 4.ii. 'About This Result' feature
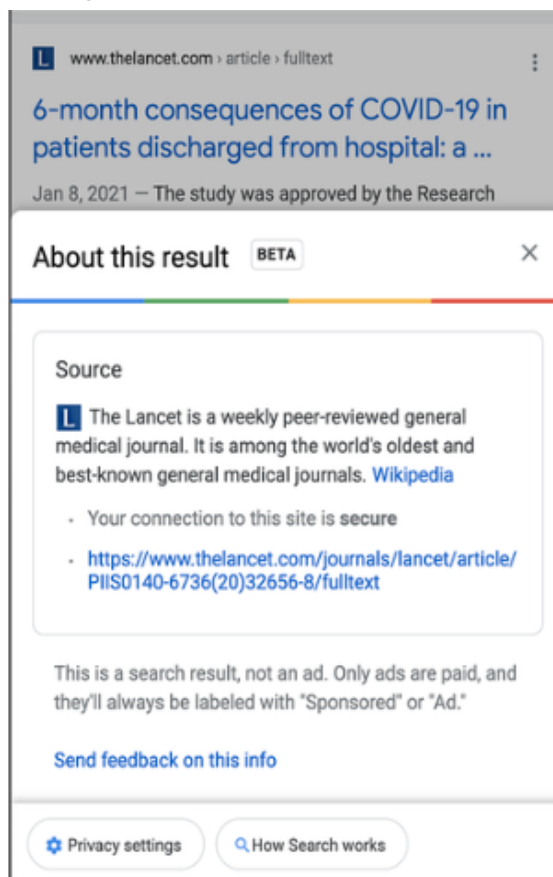
Fact checks on Google Search and News

Fact checking is an important part of tackling misinformation. As such, we have developed policies and processes on Google Search to label fact-checked articles that are displayed in Search results. When Search returns results that include fact-checked articles, we will label them as such (if they meet our eligibility and technical criteria). Google also provides tools like Fact Check Explorer and the Google FactCheck Claim Search API. The 'Fact Check' label in Google Search applies to published stories with fact-checked content that is indicated by the schema.org ClaimReview markup, like round-up stories that contain multiple fact check analyses within a single article. Google News may apply this label to publisher content, where applicable. This helps users find fact-checked content for major stories appearing on Google News.

YouTube recommendation systems

YouTube's recommendation system aims to reduce recommendations of borderline content i.e. material that comes close to, but does not cross the line of violating our policies including content that may misinform users in harmful ways, such as videos promoting a fake miracle cure for an illness, or videos claiming that the earth is flat. We continue to refine our systems where we operate, including in Australia, to address the evolving nature of misinformation. In 2021, YouTube published a blog post, 'On YouTube's Recommendation System,' which provides a detailed explanation of how our recommendation system works, including how we identify harmful misinformation or borderline content.

Health Information Panels on YouTube

As part of YouTube's investment to grow the platform's role as an effective, engaging and trusted tool for public health communication, we have expanded our partnerships with leading health organisations. We have added health source information panels to help viewers identify videos from authoritative sources and health content shelves that more effectively highlight videos from these sources when searching for specific health topics. These context cues are aimed at helping people more easily navigate and evaluate authoritative health information. To identify the right sources to include in these new features, we applied the principles developed by an expert panel convened by the National Academy of Medicine. The World Health Organization and National Academy of Medicine recently convened a meeting of interdisciplinary experts from around the world to review and validate these principles for global application. While only accredited health organisations and government entities are currently included in our health context features, we are exploring ways to broaden eligibility and evaluate inclusion of other health sources, as well as ways to expand these features globally. Additional information on our health source information panels can be found here.

Product policy updates in response to COVID-19 misinformation

In light of misinformation risks associated with the COVID-19 pandemic, Google continues to re-evaluate and update policies to help users make informed choices about COVID-19 related issues, such as vaccines and COVID-19 related products (e.g., false cures). A number of updates were introduced throughout 2021, including:

- A new comprehensive experience on Google Search related to COVID-19 that provided easy access to authoritative information from health authorities alongside new data and visualisations. This format organised the search results page to help people easily navigate to trusted resources and made it possible to add more information as it became available over time. As part of this, Search also introduced drug-specific information features for the Pfizer, Moderna and Astra-Zeneca vaccines;
- Updates to permissible mask types that may be listed in Google Shopping Ads or free listings;
- Updates to approved COVID-19 related products eligible on Google Shopping Ads (e.g. COVID-19 test kits); and
- Expanding access to personalised advertising related to COVID-19 issues for recognised and official government organisations on Google Ads.

Working closely with leading health experts, Google launched a centralised hub that includes authoritative information and insights about COVID-19, as well as support resources and tips for preventing the spread of the virus. Moreover, to help elevate authoritative information related to COVID-19, Google launched the Google Ad Grants Crisis Relief program. Our Ad Grants Crisis Relief program is awarding in-kind Ad Grants to help the World Health Organization (WHO) and global government entities provide ads on the Google.com search result pages to direct users to authoritative information regarding COVID-19.

A detailed list of Google Ads policy updates pertaining to COVID-19 can be found on this page.

Objective 5: Improve public awareness of the source of political advertising carried on digital platforms.

Outcome 5: Users are better informed about the source of Political Advertising. Signatories detail policies that provide users with information about the source of Political Advertising carried on digital platforms.

Google's broader ads policies, as described in Outcome 1b above, apply to all ads, including election ads. Election ads in Australia are ads that feature: an Australian federal political party, a candidate for the Australian House of Representatives or Senate, or a current elected federal office holder in the Australia House of Representatives or Senate. Specific measures relating to federal election ads include:
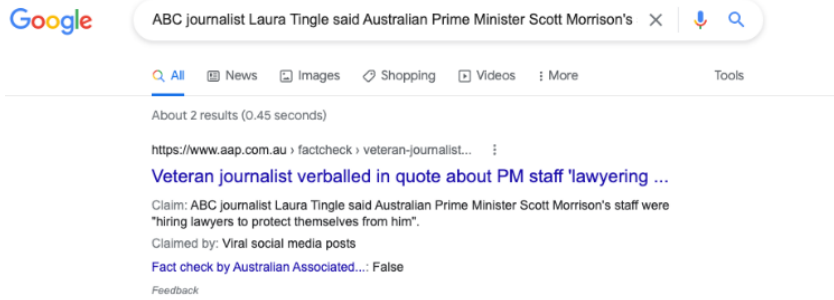
- Requirement for all advertisers running election advertising in Australia to complete the Google verification process (additional details on what constitutes an election ad in Australia and the advertiser verification process is provided in *Appendix C*);
- In-ad 'paid for by' disclosures to inform users of the source of election ads on Google services;
- Restricted criteria for election ads targeting;
- A publicly accessible transparency report webpage focused on Australian election ads, including ads spend metrics, downloadable datasets, and accompanying FAQs.

We believe that our election ads transparency tools are a vital way to provide greater protections to internet users and add important accountability on online advertisers.

Objective 6: Strengthen public understanding of Disinformation and Misinformation through support of strategic research.

Outcome 6: Signatories support the efforts of independent researchers to improve public understanding of Disinformation and Misinformation.

We continue to support global and Australia-specific efforts to enhance awareness and detection of misinformation and promote authoritative sources of information. Examples of these collaborations are provided in the table below.

Name of university, institute, or company	Overview of activities
Australian Associated Press (AAP)	AAP is providing translated fact checks to approximately 40 culturally and linguistically diverse news publishers around Australia. Collaborating with the Special Broadcasting Service (SBS), they are translated into Mandarin, Arabic and Vietnamese. The number of partners will likely increase and currently includes titles that serve multicultural audiences including: Koori Mail, Indian Link, Viet News, Epoch Times, Ngaarda Radio, Korean Herald, Australian Muslim Times, Phil Times, Almestaqbal and SBS's channels and platforms. AAP is a verified fact checker through Google, so their pieces also appear in Google search results (see below).
First Draft	First Draft is working with at least 10 news publishers (a mix of small and large) and additional community organisations to help them identify false, misleading and confusing claims during the upcoming election period. They support journalists through simulations and

	masterclasses and create an online community for the partner publishers. First Draft is also providing daily alerts and weekly briefings on false and misleading claims that are circulating online. The partners they will work with include: Asian Alliance Australia, AAP, Canberra Times, SBS, National Indigenous Television, Launceston Examiner, Northern Daily Leader Tamworth Wimmera Mail-Times, Codebreakers, Democracy in Colour, Crikey, Joy, Guardian Australia, 3CR Community Radio, Royal Melbourne Institute of Technology (RMIT) ABC Fact Check, The Drum, ABC and, The Humanism Project.

Objective 7: Signatories will publicise the measures they take to combat Disinformation.

Outcome 7: The public can access information about the measures Signatories have taken to combat Disinformation and Misinformation.

In Objective 1 / Outcome 1d of this report, we provide details (and links to corresponding materials) regarding publicly available and accessible information outlining measures we take to combat misinformation.

Additionally, as part of the Inside YouTube section of the YouTube Official Blog, we shared our approach and commitment to mitigating misinformation on the platform: increase the information from trusted sources and reduce the spread of harmful misinformation, while continuing to protect free speech and diversity of opinions. Through this section of the Blog, we continue to provide updates on our efforts to tackle this important challenge.

Concluding remarks

Through the framework of the Objectives and Outcomes set forth in the Australian Code of Practice on Disinformation and Misinformation, we have discussed how our products and services aim to mitigate the risk of harms arising from misinformation. We have outlined the various policies, safeguards and measures implemented across our products and services during 2021 to reduce the propagation and impact of misinformation. The case studies, examples and quantitative data points provided highlight the impact of these approaches. Additionally, we have detailed how we engage with the public and users, how we support other organisations tackling this issue, and how we provide further transparency regarding our own ongoing efforts.

We recognise that misinformation, and the risks associated with this issue, will likely continue to evolve. As such, we will continue to evaluate and adapt the measures and policies that we put in place across our products and services and invest in developing effective ways to protect our users and the integrity of our services. As we continue to report on the progress of this work, we look forward to engaging further with DIGI, the industry, civil society, users, academia and other key stakeholders on this issue.

Appendices

Appendix A: Google Misinformation/Disinformation Policies

Google Search & Google News Policies

- **Google Search Webmaster Guidelines**: Our webmaster guidelines prohibit techniques which could be used to deceive our ranking systems or abuse our users.

- **Google Search Content Policies**: These policies apply to content surfaced anywhere within Google Search, which includes web results. These policies include additional Search Features policies that apply to Google News, which include but are not limited to:
    - **Deceptive practices policy:** This policy prohibits sites or accounts that impersonate any person or organisation, or that misrepresent or conceal their ownership or primary purpose. We do not allow sites or accounts that engage in inauthentic or coordinated behaviour that misleads users.
    - **Manipulated media policy:** This policy prohibits audio, video, or image content that's been manipulated to deceive, defraud, or mislead by means of creating a representation of actions or events that verifiably didn't take place.
    - **Medical content policy:** This policy does not allow content that contradicts or runs contrary to scientific or medical consensus and evidence-based best practices.

- **Google News Policies:** These content and behaviour policies help ensure a positive experience for users and publishing partners. Along with Google Search's overall Content Policies, Google News has additional feature-specific policies as noted below:
    - **Ads & sponsored content policy:** Ads and other paid promotional material should not exceed content on pages. This policy states that we do not allow content that conceals or misrepresents sponsored content as independent, editorial content.
    - **Misleading content policy:** This policy states that News does not allow preview content that misleads users to engage with it by promising details which are not reflected in the underlying content.
    - **Transparency policy**: This Google News policy notes that news sources on Google should provide clear dates and bylines, as well as information about authors, the publication, the publisher, company or network behind it, and contact information.

Google Ads Policies

- **Google Ads Policies:** These policies are designed not only to abide by laws but also to ensure a safe and positive experience for our users. This means that our policies prohibit some content that we believe to be harmful to users and the overall advertising ecosystem. Our policies cover four broad areas:
    - Prohibited content
    - Prohibited practices
    - Restricted content and features
    - Editorial and technical.

- **Google Ads Misrepresentation policy**: This policy prohibits content and behaviours that deceive users by excluding relevant product information or providing misleading information about products, services, or businesses. Violations of this policy, although not a comprehensive list, include unacceptable business practices, coordinated deceptive practices, unreliable claims (such as making claims that are demonstrably false and could significantly undermine participation or trust in an electoral or democratic process) and misleading representation.
  - **Unacceptable business practices policy** does not allow:
    - Scamming users by concealing or misstating information about the advertiser's business, product, or service;
    - Ad destinations that use 'phishing' techniques to gather user information.
  - **Coordinated Deceptive Practices policy** prohibits two practices:
    - Coordinating with other sites or accounts and concealing or misrepresenting your identity or other material details about yourself, where your content relates to politics, social issues, or matters of public discern;
    - Directing content about politics, social issues, or matters of public concern to users in a country other than your own, if you misrepresent or conceal your country of origin, or other material details about yourself.
  - **Misleading representation policy** prohibits advertisers from:
    - Making misleading statements, obscuring, or omitting material information about your identity, affiliations or qualifications;
    - Providing an inaccurate business name or business name that does not clearly represent the advertised business or disambiguates from similar businesses in the ad or user interactions.
  - **Manipulated media policy** does not allow advertisers to manipulate media to deceive, defraud, or mislead others.
  - **Unreliable claims policy** does not allow advertisers to make inaccurate claims or claims that entice the user with an improbable result as the likely outcome a user can expect. This includes claims related to health and weight loss, financial products or money making schemes, and politics, social issues, or matters of public concern.

- **Google Ads Inappropriate content policy**: As noted in our help centre, this policy prohibits ads or destinations that display shocking content or promote hatred, intolerance, discrimination, or violence.

Google AdSense Policies

- **Google Publisher policies**: Users who monetise content with Google must adhere to all Google publisher policies. Prohibited content and behaviours include but are not limited to: dangerous or derogatory content, misrepresentative content (such as making claims that are demonstrably false and could significantly undermine participation or trust in an electoral or democratic process) or coordinated deceptive practices.

- **AdSense Program policies**: AdSense publishers are required to adhere to the AdSense Program Policies or risk their ad or account being disabled. It is the publisher's responsibility to keep up to date with and adhere to the following policies: invalid clicks and impressions,

encouraging clicks or views (non-rewarded inventory), traffic sources, ad behaviour, ad placement, site behaviour, deceptive site navigation.

YouTube Policies

- **YouTube Community Guidelines:** These Community Guidelines outline what type of content is and is not allowed on YouTube. These policies apply to all types of content on our platform, including videos, comments, and links. These policies cover: spam & deceptive practices, sensitive content, violent or dangerous content, regulated goods, misinformation and monetisation.

- YouTube Misinformation policies detailed below can be found on this landing page.
  - **YouTube misinformation policies:** Certain types of misleading or deceptive content with serious risk of egregious harm are not allowed on YouTube. This includes certain types of misinformation that can cause real-world harm, like promoting harmful remedies or treatments, certain types of technically manipulated content, or content interfering with democratic processes.
  - **YouTube election misinformation policies:** Certain types of misleading or deceptive content with serious risk of egregious harm are not allowed on YouTube. This includes misinformation that can cause real-world harm, like certain types of technically manipulated content, and content interfering with free and fair democratic election processes.
  - **YouTube COVID-19 medical misinformation policy:** YouTube does not allow content that spreads medical misinformation that contradicts local health authorities' (LHA) or the World Health Organization's (WHO) medical information about COVID-19. This is limited to content that contradicts guidance on treatment, prevention, diagnosis, transmission, and the existence of COVID-19. Note: YouTube's policies on COVID-19 are subject to change in response to changes to global or local health authorities' guidance on the virus. There may be a delay between new LHA/WHO guidance and policy updates given the frequency with which this guidance changes, and our policies may not cover all LHA/WHO guidance related to COVID-19.
  - **YouTube vaccine misinformation policy:** YouTube does not allow content that poses a serious risk of egregious harm by spreading medical misinformation about currently administered vaccines that are approved and confirmed to be safe and effective by LHA and the WHO. This is limited to content that contradicts LHA or WHO guidance on vaccine safety, efficacy, and ingredients.

- **YouTube Impersonation policy**: This policy states that content intended to impersonate a person or channel is not allowed on YouTube. YouTube also enforces trademark holder rights. When a channel, or content in the channel, causes confusion about the source of goods and services advertised, it may not be allowed.

- **YouTube fake engagement policy**: YouTube does not allow anything that artificially increases the number of views, likes, comments, or other metrics either by using automatic systems or serving up videos to unsuspecting viewers. Content and channels that do not follow this policy may be terminated and removed from YouTube.

- **YouTube spam, deceptive practices, and scam policies**: YouTube does not allow spam, scams, or other deceptive practices that take advantage of the YouTube community. We also do not allow content where the main purpose is to trick users into leaving YouTube for another site.

- **YouTube advertiser-friendly content guidelines**: Users in the YouTube Partner Program can share revenue from ads. This policy exemplifies content that is not suitable for ads that will result in a 'limited or no ads' monetisation state.

- **YouTube channel monetisation policy**: YouTube monetisation policy includes YouTube's Community Guidelines, Terms of Service, Copyright, and Google AdSense program policies. YouTube enforces this monetisation policy by:
  - Turning off Ads from your content;
  - Suspending your participation in the YouTube Partner Program; and
  - Suspending or even terminating your YouTube channel.

Appendix B: Tools and features used to empower users in making informed choices

Google has developed many features and interventions to provide more context to users and ensure that authoritative sources are elevated in response to searches or browsing intents related to health, civic participation, current events, and other topics where users want content that they can trust. The features and interventions listed below are all available to Australian users of our services.

- **Surfacing Fact-Checks on Google Search, Images, and News**: easily enable users to find articles that fact check public information;

- **Elevating original reporting in Google Search**: provide greater context to subsequent news stories;

- **'Full Coverage' in Google News**: help users access context and diverse perspectives about the news stories they read;

- **'Information Panels', including Fact Check, and those giving publisher context and topical context on YouTube**: provide greater context around topics related to searches or videos (additional information can be found here);

- **'Breaking News' and 'Top News Shelves' on YouTube**: prominently surface news content from authoritative sources only;

- **Authoritativeness in YouTube recommendations**: prioritise information from authoritative sources for newsworthy events and topics prone to misinformation in search results and recommendations (additional information can be found here);

- **Labelling state-funded news channels on YouTube**: label publishers that are government or publicly funded on the watch pages of their videos.

Appendix C: Advertiser Verification Process

Election ads in Australia are ads that feature an Australian federal political party, a candidate for the Australian House of Representatives or Senate, or a current elected federal officeholder in the Australia House of Representatives or Senate. Additional requirements must be met by advertisers who want to run election ads in Australia, including advertiser verification by Google. More information can be found here.

In 2020, Google introduced a new advertiser verification program to capture more information from all our advertisers about their identities and businesses. The program was phased in on a rolling basis geographically; rollout in Australia began in Q1 2021. The program verifies identification and then discloses the advertiser behind specific ads in the 'Why this Ad' (in some places, 'About this Ad') menu. This verification and disclosure feature applies to the ads that Google shows on our own properties and those of our publishing partners. Note that the disclosures view can vary slightly between products.

See screenshot below:
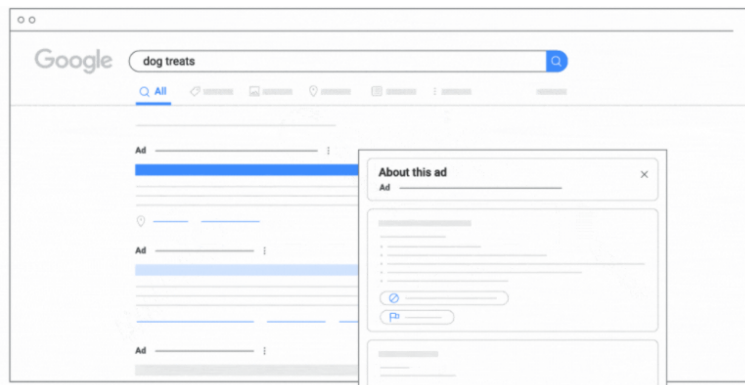
Verification process for Australian advertisers
Prospective advertisers who wish to run election ads in Australia are also required to go through a verification process. Once Google verifies the advertiser's eligibility to run election ads, they receive an email and an in-account notification. Verifying their identity may require two steps and each step can take up to 5 business days. Our teams are trained to handle this process at scale across Australia, and are equipped to respond to related questions from the political parties and candidates participating in, and institutions responsible for, Australian elections.