Australian Government

Digital Transformation Agency

Understanding the common pain points using web analytics

June 2021

dta

Digital Transformation Agency

Version: 1801

Contents

Glossary

Term	Definition
Bounce rate	The percentage of users who leave a site after viewing a single page without triggering an event.
Direct traffic	A session that is initiated by a user typing the url of a website into the browser or using a bookmark.
Event	Recorded user interactions with a website. For example, a scroll depth event at 50% occurs when a user scrolls to 50% on a page where the scroll depth event has been set up to be tracked. The actions that trigger an event being recorded are determined by event configuration for individual websites.
Landing page	The first page a user visits when accessing the website.
Organic search	Search results that are not paid advertisements.
Pageview	An instance of a user visiting a particular page on a website.
Referral	A user accessing a website using a link on another website.
Session	A group of user interactions with your website that take place within a given time frame by an internet user.
Traffic	A collective way of referring to users visiting a website. Traffic is measured in sessions.
Unsampled Data	Unsampled data includes all collected data points. The Google Analytics interface shows a sample of the collected data points, to represent the data while reducing the complexity of processing it. Unsampled Google Analytics data can be extracted from BigQuery.
User	A person that uses a digital device to visit a website.

In depth definitions of metrics used in this analysis are provided at Appendix B.

1 Executive Summary

This research explores the common pain point 'Difficult to find information' using web analytics data. The Digital Transformation Agency's (DTA) GovX and Observatory teams collaborated on this research, using the Observatory's data and analysis techniques to create new perspectives on the common pain points identified by GovX. The research focused on 'Difficult to find information' in the context of Australian Government websites at the federal level. Key findings from our research include:

- Returning users tend to visit more pages and spend more time on a page, suggesting that they engage with content differently
- People accessing websites through an external search engine seem to find the content they are looking for more directly than other users, viewing fewer pages per session and spending less time on page
- Referrals from other Australian Government websites are a significant source of traffic, indicating that some people need to visit multiple government sites to find what they are looking for.

Patterns in web analytics can show where to target research using other methods to understand what is happening. Evaluating web analytics in the context of website's functions and purpose is a powerful way to find opportunities for policy and product owners to improve people's experience.

This project developed benchmarks for five key metrics so agencies can compare their sites to the broader Australian Government. We also created a Google Data Studio tool that makes it simpler for agencies using web analytics to diagnose this common pain point and identify areas of their site where further research would be beneficial.

Identifying if a pain point related to their website is an instance of a common pain point allows agencies to explore existing solutions from other agencies. Over time, it may identify opportunities to build scalable, collaborative solutions that mitigate the problem across both Commonwealth and State and Territory governments. These solutions may require changes to technology, processes or policy.

2 Background

2.1 The project

This project was a collaboration between analysts from two teams at DTA:

- The **GovX team** conducts research and analysis to drive change across government to improve people's experience during key life events.

- The **gov.au Observatory** informs decision makers, policy development and service improvement through data-driven insights.

GovX identified 11 common pain points to help government better understand the pains people experience when accessing products and services. The common pain points (provided at Appendix A) were identified by analysing pain points that occur repeatedly across multiple life events. They were first documented in 2018 and were most recently updated in January 2021 to reflect recent life event research.

After the 2021 update, GovX approached the Observatory to develop a collaborative research project. The project combines their complimentary research methodologies to explore a common pain point using the Observatory's web analytics data and GovX quantitative analysis techniques. The collaborative research project had two broad goals:

1. Explore how web analytics can contribute to understanding a specific common pain point across the Australian Government at the federal level

2. Develop ways to support agencies to use their own web analytics data to explore the common pain points in their own context

Initial scoping activities identified the common pain point 'Difficult to find information' (described in section 2.2 below) as a focus for the exploratory analysis. The team developed accompanying comparative benchmarks (discussed further in sections 3 and 4.2) to enable policy and product owners to contextualise their own data against

this research. Policy and product owners can use these benchmarks to compare key metrics from their site to the broader Australian Government. This type of comparison has not been previously possible, as agencies only have access to their own data.

While web analytics are a powerful way to identify potentially problematic areas, they are most effective when used contextually. The same metrics can mean different things in different contexts (discussed further in section 3), and this context needs to be understood to interpret metrics accurately. The team developed a Google Data Studio template so agencies using Google Analytics can apply this research directly to their own data. Agencies can use this tool to compare their data to the government benchmarks and explore additional indictors that information may be difficult to find. This tool will help product and policy owners to use web analytics to interrogate this common pain point in their own context.

Web analytics are an effective way to detect patterns and anomalies, making them useful for identifying where a common pain point may be occurring. However, it is important not to extrapolate beyond the available data. Web analytics should be used in conjunction with other research methods to investigate whether the indicator is accurately detecting a problem. Web analytics show us *where* patterns occur, but targeted research is required to understand *why* they occur.

2.2 Difficult to find information

Difficult to find information

People experience challenges finding the information they need. This could be because it is spread across multiple sources or the amount of information is overwhelming.

'Difficult to find information' is one of the 11 common pain points that people experience when interacting with government. GovX identified this issue in 9 of the 11 life event journeys analysed for the common pain points research:

- Accessing childcare
- Australians travelling overseas
- Coming to Australia to study
- Experiencing a natural disaster
- Having a baby
- Importing goods
- Looking for work impacted by COVID-19
- Transitioning to retirement
- Youth transitioning from education to employment

The GovX life event research found that people can be negatively affected by being unable to find information, or by receiving inconsistent or inaccurate information when interacting with government. People are often unsure whether they have found the right information when there is a lot of information, or when it is spread across multiple sources. Difficulty finding information is made worse during crisis situations when information is often changing rapidly. People described 'going in circles' or needing to 'piece things together' from multiple sources when attempting to find information, causing frustration and uncertainty.

"I found *[finding information online]* a little convoluted actually, I knew what I was looking for so I think that was helpful, but I think if you didn't you tend to find yourself getting thrown in a bit of a circle looking for things. Sometimes it's picking the right word when doing your search." **User interview, Transitioning to retirement**

3 Method

3.1 Available data

The Observatory administers the Australian Government google analytics subscription which collects anonymous user interaction data. The research team used the Observatory's unsampled Google Analytics 360 (GA360) dataset for this analysis as it is the most comprehensive data available. Additionally, Observatory subscriber agencies can access comparable data, so they can apply the research to their policies and services.

The research data comes from 1 January to 31 March 2021 (inclusive). This period sits outside the coronavirus traffic spike that government services experienced in 2020. The dataset only covers Observatory subscriber agencies at the time, and includes 39 Commonwealth agencies, entities, or companies. This constraint in the research dataset should be considered when applying the findings to your policy and services. This research examines 'Difficult to find information' in the context of Australian Government websites, although the broader common pain points research looked at interactions with all levels of government, as well as non-government services.

3.2 Selected metrics

The team chose five key metrics for analysis and benchmark development. Three of the metrics are provided directly by GA360, and two are calculated from other dimensions in the data. The selected metrics were:

- Average pages per session (GA360)
- Average time on page (GA360)
- Bounce rate (GA360)
- Proportion of traffic generated by new users (calculated)
- Proportion of government referrals (calculated)

It's important to consider how each metric is defined and any related limitations when interpreting these metrics. For example, a "user" does not equate directly to an individual person; and a "new user" is likely to be someone who has not visited the website before, but this is not always the case. Average pages per session, average time on page, and bounce rate are all metrics that are often used as proxies for user engagement. It's important to examine the assumptions associated with how these figures are judged. A user visiting a high number of pages could be having difficulty finding information; while someone visiting only one or two pages may have been able to find what they need quickly. Full data definitions and discussion of constraints and considerations associated with each metric are provided at Appendix B.

3.3 Analysis method

The team produced a series of aggregated tables (by date, domain, and domain category) in Google BigQuery SQL. The SQL code used is available on the Observatory's Github page.

We used the 'by date', 'by domain category', and 'by region' aggregation tables to conduct the exploratory analysis. The analytical techniques used to detect patterns and anomalies within the data include:

- examination of descriptive statistics
- visual analysis of data distributions
- time series analysis
- comparative cohort analysis.

We looked at breakdown dimensions such as user type, traffic source, session timing, user location, and website category (listed in Appendix B.11). Findings from the exploratory analysis are in section 4.1.

The benchmarks are based on the 'by domain' aggregation table, after removing outlying datapoints. Each benchmark is published as a collection of descriptive statistics and an accompanying histogram. The benchmarks and discussion of considerations for how to use them effectively are included in section 4.2.

4 Results and discussion

4.1 Exploratory research findings

The most interesting insights our analysis revealed were detected through comparative analysis between different types of users. For example:

- new and returning users behave differently.

- Users accessing the site via a search engine have a different engagement pattern to traffic from other sources.

While the patterns we found are relevant to the common pain point "Difficult to find information", it must be reiterated that web analytics can detect patterns but not explain them. The findings are presented alongside a series of key questions which can be used to guide further contextual user research.

4.1.1 Returning users are more engaged

Returning users are generally more engaged on websites than new users. They typically spend distinctly more time on pages, visit more pages per session, and have a lower bounce rate. This means that returning users spend more time interacting with more content and are less likely to immediately exit a website.

Key statistics

- Returning users on average spent 477 seconds on a webpage across the subscribed government websites compared to new users who spent closer to 220 seconds on a webpage, less than half that of returning users

- Returning users looked at an average of just over 15 pages per session compared to new users looking at closer to 10 pages per session

- The average bounce rate for returning users was around 15%, compared to 25% for new users.

Key questions

- Why do people behave differently on return visits? Do they have different needs or goals?

- Do return users spend more time on the website and engage with more content because they are doing more complex tasks than new users?

- Are new users more likely to have difficulty finding information because they are less familiar with a website? If so, what changes can be made to the website to assist new users?

- Are new users more likely to bounce because they decide they won't find the information they need on the website?

4.1.2 Usage is higher on weekdays than weekends

The daily traffic on government websites is higher during the weekdays than on weekends. Weekdays have a lower bounce rate than weekends, and users are more likely to be returning users. While the day of the week doesn't directly affect whether information is difficult to find, users visiting for different purposes are likely to have different experiences when trying to find information.

Key statistics

- The average number of users on weekend days was 52% less than on weekdays

- The peak usage in the week over all government websites was on Wednesdays

- 56% of weekend traffic was generated by new users, compared to an average of 45% on weekdays

- The bounce rate for returning users was about 12% on weekdays and closer to 17% on weekends

- There was not distinct difference in the time on page or the pages per sessions on the weekends.

Key questions

- Do weekday and weekend users have different goals and needs?

- Are people accessing websites on weekdays more likely to be doing something related to their work or studies?

- Are weekend users more likely to be completing personal tasks?

- Are weekday users more likely to regularly access the same content?

- Are weekend users more likely to need to complete 'one-off' tasks?

4.1.3 Organic traffic finds content most directly

People accessing websites through an external search engine (categorised by google as 'organic traffic') view the fewest pages per session and had the lowest average time on page. It's likely that users accessing websites through external search find the content they are looking for more directly, and do not need to engage with as much content when looking for information. Organic search is a significant traffic source for government websites, with over half of all sessions initiated via a search engine.

Key statistics

- Organic searches were the source of 51.7% of traffic

- Organic search traffic had an average of nine pages per session, compared to 15 pages for referral traffic and 20 pages for direct

- The average time on page for organic search traffic was 316 seconds, which is 42 seconds less than the overall average.

Key questions

- Do users accessing the site from a search engine land on expected or unexpected pages?

- Are people who access the site using a search engine able to reach their goal more quickly?

- What sort of content are people engaging with? Are high or low pages per sessions expected for this site?

- Are lower pages per session an accurate indicator that people can find information more quickly on this site?

4.1.4 Australian Government referrals are the largest traffic referrer

Other Australian Government websites were the dominant source of referral traffic, and a significant source of overall traffic. High government referrals are likely to be an indicator that information is difficult to find, with the GovX research underpinning the common pain points identifying that fragmentation of content across multiple government websites can cause frustration and uncertainty. This fragmentation may cause users difficulty even if they are able to find information quickly once they reach the correct website. The proportion of referral traffic attributed to government websites varies considerably between different website categories, with business and employment having the highest rate of government referrals.

Key statistics

- The Business and Employment category of government domains had a large percentage of Australian Government referrals at over 80%

- Australian Government websites are the source of 72% of referral traffic, and Australian Government referrals make up 18% of all traffic

- Out of the websites that refer traffic to government websites, 27% were non-government websites, 68% were Australian Government sites and 5% were state and local government sites.

Key questions

- What sort of content are people accessing the site via another site visiting?

- Which other government sites are referring traffic to this site?

- Where does this site refer traffic to?

- Are the referral pathways expected or surprising? Is there a way they could be optimised?

4.1.5 Australian Government referral traffic engages differently

Traffic that was referred from another Australian Government website behaved significantly differently to traffic from other sources. Australian Government website referrals had the highest average time on page, second highest pages per session, and the lowest bounce rate. While this may indicate high levels of engagement, users referred from other government websites could also be experiencing difficulty finding information (discussed in section 4.1.4).

Key statistics

- Pages per session were very high for Australian Government website referrals, with traffic from this source viewing an average 18 pages per session. Only direct traffic was higher, with an average of 20 pages per session.

- The average time on page for Australian Government website referrals was 471 seconds, which is nearly two full minutes longer than the overall average of 358 seconds.

- Australian Government website referrals had the lowest bounce rate with only 8.6% of sessions bouncing, over 10 percentage points lower than the average.

Key questions

- Are users referred from other government websites having more difficulty finding the information they need?

- Are users who come from another government website completing more complex tasks?

4.1.6 The ACT has significant traffic per capita

The ACT had disproportionally high traffic per capita compared to the other jurisdictions. Traffic from the ACT also behaved differently, with the highest average time on page, lowest bounce rate, and lowest proportion of traffic generated by new users. It's possible that this anomaly is a result of the high proportion of public servants in the ACT, with this workforce being likely to more frequently access government websites as part of their duties. Although this is not directly linked to the common pain point 'Difficult to find information' (this workforce was not a focus of the underlying life events research), it may be useful in identifying a key user cohort.

Key statistics

- The ACT's per capita traffic to government websites was nearly three times higher than that of other jurisdictions. The ACT had an average daily sessions

rate of 43 per one hundred people, while other jurisdictions ranged from 5 (Northern Territory) to 17 (New South Wales).

- The ACT had the lowest proportion of traffic generated by new users, at around 41%. The average across all jurisdictions (including the ACT) was 45%.

- The ACT had the highest average time on page, at 420 seconds, over a minute longer than the overall average of 346 seconds.

- The bounce rate for the ACT was 14%, lowest of all jurisdictions. This was 3 percentage points lower than the overall average of 17%.

- Pages per session was reasonably consistent across jurisdictions; with the ACT averaging 12 pages per session and all other jurisdictions averaging 11.

- These results remained consistent when Canberra was compared with other metropolitan areas, such as Sydney and Melbourne. This suggests that the difference between the ACT and other jurisdictions is not because the ACT is predominantly made up of a metropolitan area.

Key questions

- Is the disproportionate traffic related to the public service workforce?

- Do public servants behave differently to other users when accessing government websites?

4.2 Government benchmarks

The team developed benchmarks for 5 key website metrics that allow agencies using GA360 to compare their sites with federal level Australian Government websites generally. The benchmarks provide context for the metrics and can be used to understand how a website sits in the ecosystem of Australian Government websites. This sort of comparison has not previously been possible for agencies, as aggregated web analytics across the Australian Government is only available in the Observatory's

data analysis environment. We also developed a Data Studio template that allows users to easily compare their data to these benchmarks. Using the benchmarks for contextual comparison is discussed in section 5.

It is important to understand that these benchmarks are not targets, goals or success metrics. There will be many reasons for a site's value for a metric to differ from the benchmark. Each of the selected metrics will indicate different things in different contexts. The benchmarks are intended to provide an additional perspective for understanding website performance, not a mechanism for assessing it. A site's position within the benchmark distribution will be influenced by the site's purpose and functionality.

4.2.1 Average pages per session

A user typically looks at 6 to 7 pages during a session when visiting a Australian Government website. Across subscribed government websites in quarter one 2021, the median pages per session was 6.59 (see Table 1 and Figure 1). Pages per session varied greatly among the subscribed government websites, as shown in Figure 1. Approximately 50% of sites had an average pages per session of between four and nine pages. The distribution's spread is probably due to the variability in function and purpose of the sites.

Table 1 Average pages per session benchmark figures

Average pages per session	
Average	7.44
Median	6.59
Standard deviation	3.67
25th quantile	4.87
75th quantile	9.05

Figure 1 Average pages per session benchmark distribution
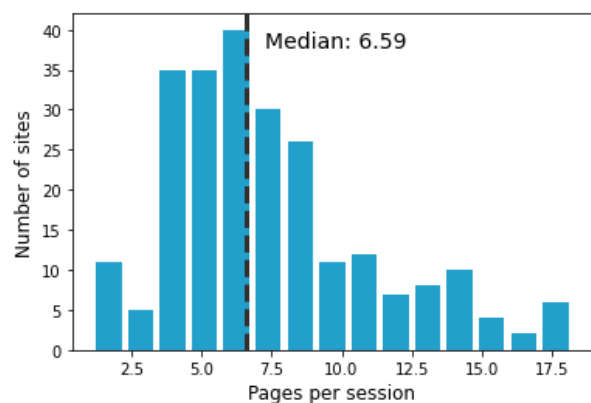
Considerations

The pages per session metric is particularly dependent on the site's purpose, functionality, and design. When applying this benchmark, it is important to consider:

- The different goals users may have when visiting the site

- The number of pages per session that would be expected for each goal

- How the site structure impacts the number of pages a user may interact with.

Some user goals may require visiting many more pages than others (for example, completing a multi-page form). Identifying groups of users using the audience segmentation functionality provided in Google Analytics can be a helpful way to detect users with different goals and conduct more targeted analysis.

A large difference between the expected and actual pages per session may be an indicator that information on your website is difficult to find.
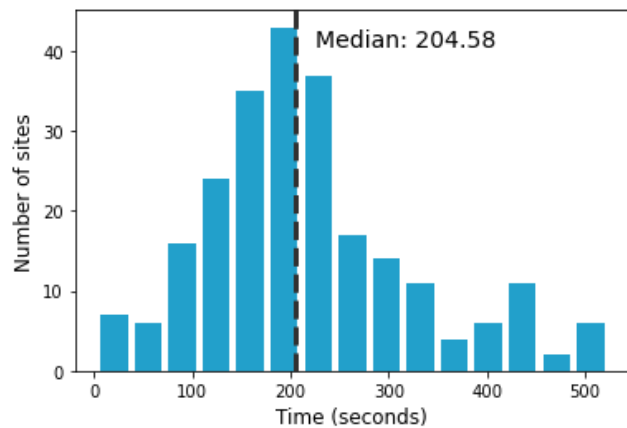
4.2.2 Average time on page

In quarter one of 2021 the average time spent viewing a page across subscribed government websites is just under 3 minutes and 30 seconds, and the median time on page was 204.58 seconds (see Table 2 and Figure 2).

Table 2 Average time on page benchmark figures

Figure 2 Average time on page benchmark distribution

Average time on page (seconds)	
Average	217.60
Median	204.58
Standard deviation	107.36
25th quantile	146.03
75th quantile	272.77

Considerations

The average time on page is an important metric that can be reflective of user engagement. However, as discussed in previous sections, it should not be assumed that time on page is directly correlated with engagement. When applying this benchmark, it is important to consider:

- How the content design and structure of the site impacts the amount of content presented on each page

- How much of the content on a page users are engaging with

- The different goals users may have when visiting the site

- The complexity of each goal and how much time it would be expected to take to achieve it.

Content structure will have significant impact on the average time on page for an individual site. Combining this metric with other indicators of user engagement and measures of content volume and complexity may provide an effective way to identify problematic areas where information is difficult to find.
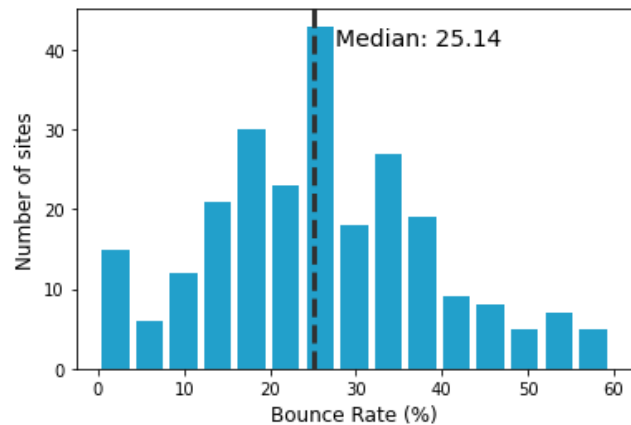
4.2.3 Bounce rate

Around a quarter of sessions started on Australian Government websites bounce, with the user not interacting further with the site after loading the first page. Across subscribed government websites in quarter one 2021, the median bounce rate was 25.1% (see Table 3 and Figure 3). Half of the subscribed sites had a bounce rate of between 17.4 and 34.1%.

Table 3 Bounce rate benchmark figures

Bounce rate (%)	
Average	25.91
Median	25.14
Standard deviation	13.16
25th quantile	17.44
75th quantile	34.07

Figure 3 Bounce rate benchmark distribution

Considerations

As discussed in section 3.1.1, although bounce rate is generally considered an indicator of user disengagement, sessions can bounce for a variety of reasons. Bounce rate can also vary depending on a site's event configuration. When applying this benchmark, it is important to consider:

- Whether a bounced session could still allow the user to achieve their goal

- The characteristics of high bounce rate sessions

- Whether any events have been configured that would affect the bounce rate

The future of bounce rate

Google Analytics 4 does not include the bounce rate metric found in Google Analytics 360. In GA4, bounce rate has effectively been replaced by the 'Engaged Sessions' metric, which defines session as engaged if the user is on the site for more than 10

seconds. This is potentially a more intuitive and accurate indicator of user engagement and will not be affected by variations based on event configuration.

The Observatory is in the initial stages of rolling out GA4 to all subscribers. There is not yet sufficient GA4 coverage to calculate a Australian Government benchmark, but future releases of the benchmarks are likely to include this metric.

4.2.4 Proportion of traffic generated by new users

The majority of subscribed sites have over 50% of their traffic generated by new users. Across subscribed government websites in quarter one 2021, the median proportion of traffic generated by new users was 64.21% (see Table 4 and Figure 4).

Table 4 Proportion of traffic generated by new users benchmark figures

Proportion of traffic generated by new users (%)	
Average	63.47
Median	64.21
Standard deviation	14.53
25th quantile	53.00
75th quantile	72.85

Figure 4 Proportion of traffic generated by new users benchmark distribution
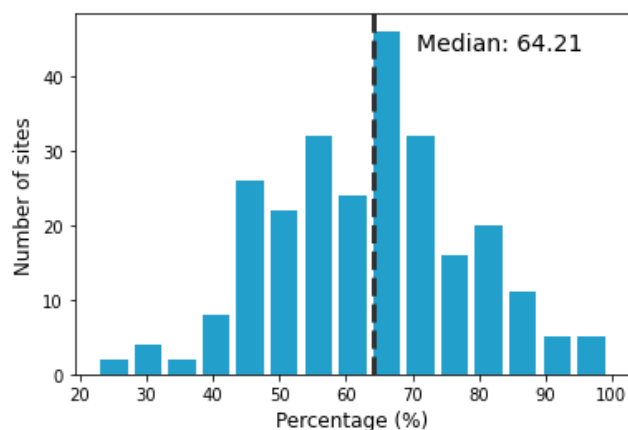
Considerations

It's important to understand that a new user does not necessarily equate to a new person visiting the website (discussed in Appendix B.5), however this metric can be used to approximate the volume of traffic associated with people who are likely not to have visited a website before.

The proportion of traffic generated by new users is particularly dependent on a site's purpose and audience. When applying this benchmark, it is important to consider:

- The different goals users may have when visiting the site

- Whether users would be expected to visit the site repeatedly

- How frequently a repeat user might need to visit the site

- Whether frequent visitors can more easily find what they need

Examining the proportion of traffic generated by new users can help to better understand a site's user base and assess how familiar users may be with the site.

4.2.5 Proportion of government referrals

Proportion of government referrals was selected as a benchmark as high rates of referral traffic from other government sites is likely to be an indicator that information is difficult to find or fragmented between government sites (as discussed in sections 4.1.4, 4.1.5, and Appendix B.8).

Across subscribed government websites in quarter one 2021, the median proportion of government referrals was 52.10% (see Table 5 and Figure 5).

Table 5 Proportion of government referrals benchmark figures

Proportion of government referrals (%)	
Average	53.37
Median	52.10
Standard deviation	27.44
25th quantile	33.92
75th quantile	75.73

Figure 5 Proportion of government referrals benchmark distribution
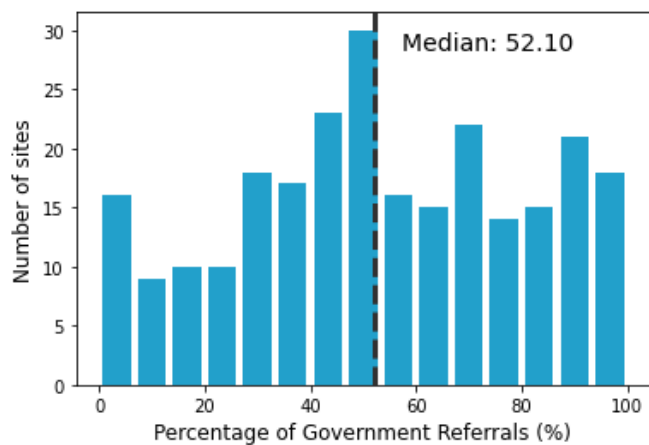
Considerations

This metric is calculated as a proportion of total referral traffic. When examining this benchmark in context, it is important to consider:

- The overall proportion of referral traffic

- Which government sites are directing traffic to the site

- What sort of content this traffic is accessing

- The goals of users referred by other government websites

- Any outbound referrals to other government sites the site is generating

User research suggests that being required to visit multiple government websites to find information is a significant cause of user pain. Understanding the proportion of referral traffic that is coming from other government websites can help to identify whether specific referral pathways are contributing to information being difficult to find.

5 Applying this research

Web analytics are most powerful when explored in context, with an understanding of the website's function and purpose. Websites with different functions will perform differently against each benchmark. Table 6 provides a matrix of suggested expectations for different types of websites.

Table 6 Expected relationship to benchmark by website type

Website type	Average time on page	Average pages per session	Bounce rate	Proportion of traffic generated by new users	Proportion of Government referrals
Benchmark median	*204 seconds*	*6-7 pages*	*25%*	*64%*	*52%*
Content based - informative	Greater than	Greater than	Less than	Less than	Greater than
Content based - blog	Greater than	Less than	Greater than	Greater than	Equal to (Depends on type of blog)
Call to action – sign up	Greater than (depends on process)	Depends on process	Less than	Greater than	Greater than
Call to Action – application / form	Greater than	Greater than	Less than	Greater than	Greater than
Contact us page	Less than	Less than	Greater than	Equal to	Less than
Online shopping	Greater than	Greater than	Less than	Less than	Less than
Promotional website	Less than	Less than	Greater than	Greater than	Greater than
Forum / discourse	Greater than	Less than	Greater than	Greater than	Equal to

This matrix is a general guide for applying the benchmarks in context. Websites, or specific areas of websites, will measure above or below the benchmark median depending on their purpose or function. For example, an informative website with dense longform content is likely to have a median time on page higher than the benchmark; while a transactional site focused on completing multi-page forms should expect a higher number of pages per session than the benchmark.

5.1 Data Studio tool

The team created a Google Data Studio template that makes it simple for anyone using Google Analytics 360 to apply this research to their own data. Users can compare their own data directly with the benchmarks, which will be updated quarterly. The tool also includes exploration of other factors relevant to 'Difficult to find information', such as a breakdown of government referral pathways or 'contact us' pageviews. Contact the Observatory to arrange access to this tool.

We created this tool to make it easier for product owners to diagnose this common pain point and identify specific areas of their site where further research would be beneficial. Web analytics are a powerful way of identifying patterns, but contextual knowledge is required to understand if a pattern indicates a problem. Further research is generally required to understand why the problem is occurring, and to find ways to solve it. Analysis of web analytics should be combined with other research methodologies. This can help to make sure that the needs and contexts of the people using the website are considered when addressing the pain they are experiencing.

6 Conclusion

This research piece provides an additional perspective on the common pain point 'Difficult to find information'. We identified patterns across the Australian Government at a federal level, which could be explored in context to establish whether they are indicators that information is difficult to find. The benchmarks and accompanying Data Studio report make it easier to conduct this exploration.

Web analytics can help us to identify manifestations of the common pain points within digital services. Identifying when their pain point is an instance of a common pain point helps agencies find opportunities to link into the wider service landscape. Agencies can investigate whether a pain point is already addressed by existing solutions they can reuse or adapt. The common pain points also show opportunities to build scalable, collaborative solutions or reusable platforms which many agencies can benefit from.

This collaboration shows how large-scale web analytics can be linked with the results of detailed qualitative research to provide holistic evidence of opportunities to improve people's experience of dealing with government.

To learn more about the **gov.au Observatory** and the services they provide, visit https://www.dta.gov.au/our-projects/govau-observatory or contact observatory@dta.gov.au. The Observatory will continue to develop Australian Government analytics benchmarks based on feedback from subscribers about what is useful.

For more information about the common pain points or life event research, contact the **GovX team** on govx@dta.gov.au or visit https://www.dta.gov.au/our-projects/govx. Previous GovX research is available on the GovX Life Event Communities. The GovX team is currently working on further guidance material on how to diagnose the common pain points in context.

Appendices

Appendix A: The Common pain points

The GovX team identified 11 common pain points to help government better understand the pains people experience. They are based on patterns of pain points that occur repeatedly across multiple life events. The common pain points were most recently iterated in January 2021. For more information about the common pain points, visit the DTA website, or contact govx@dta.gov.au or join the GovX life event communities at https://lifeevents.digital.gov.au

The eleven common pain points are:

1. Not aware of progress or important changes

People don't know about where processes are up to, why things happen or important changes. This could be because they did not receive or access information. It becomes difficult for them to plan and can put them into an avoidable negative situation.

2. Difficult to find information

People experience challenges finding the information they need. This could be because it is spread across multiple sources or the amount of information is overwhelming.

3. Difficult to understand or use information

People can find it difficult to make sense of or use information. This can be because the language is too complex, it is not available in their preferred language, or is presented in a format they can't access. Sometimes information can be inaccurate or out of date.

4. Not aware of available support

People are not aware of services and resources that could support them. This means that they do not look for or access support or information that could improve their situation.

5. Uncertain about entitlements and obligations

People are unsure about their entitlements and obligations when interacting with service providers. The policies, rules and processes are often complex and difficult to understand. This means that they can miss out on support or not fulfill obligations.

6. Unclear how to reach an outcome

People are uncertain about what they need to do to reach an outcome, particularly when they need to interact with multiple services or organisations. They often feel lost and frustrated and may need to ask for help.

7. Too much time and effort to complete tasks

Accessing services can be time consuming and difficult. This can include waiting in person or on the phone, putting together applications, gathering evidence, following up on mistakes or checking progress. Long times taken to receive services can cause challenges when people need immediate support.

8. Required to provide information multiple times

People find it frustrating to provide the same information multiple times. This is magnified when they have already provided the information as part of achieving the same or a similar outcome.

9. Unable to use preferred channel

People are not able to use their preferred channel to complete a task. This could be because they are pushed towards a particular channel, experience issues using a channel or need to interact with multiple channels. As a result, tasks are more difficult or impossible for them to complete without extra assistance.

10. Difficult or impossible to complete task

People face challenges using services that do not account for their needs and circumstances. Needs and circumstances can be diverse and influenced by factors such as age, gender, culture, language, ability, location, or socio-economic status. This means people may not be able to use a service, use work arounds or settle for an unsuitable interaction.

11. Not getting the right support

People feel that they aren't getting the right support to help them reach an outcome. This can be for a range of reasons including the support not being enough, the staff they deal with not having the right skills, a lack of trust in the organisations they have to deal with and or a perception of poor service quality.

Appendix B: Data definitions, constraints, and considerations

B.1 Average pages per session

Definition: The average number of pages users visit during a session. Pages per session is calculated as:

$$\frac{(Total\ pageviews)}{(Total\ sessions)}$$

Considerations: Pages per session is often used as a proxy for user engagement, with increased pages taken to mean higher engagement. This should be carefully considered in context, as a user visiting a high number of pages could be having difficulty finding the information they are looking for, while a user visiting only one or two pages may have been able to directly access the information they needed.

B.2 Average time on page

Definition: The average of the recorded 'time on page' measurements. Time on page is calculated as the duration between when a user loads a page and when they load the next page they visit, measured in seconds.

The average time on page is calculated as:

$$\frac{(Total\ time\ on\ page)}{(Total\ pageviews) - (Exit\ pageviews)}$$

Constraints: The time on page calculation is affected when the page is the last page a user visits (the 'exit' page), as there is no subsequent page load. An exit page will have a time on page measurement of 0, unless the user triggers an interactive event, in which case the time on page will be the time between the page load and the final interactive event.

Considerations: Average time on page will be slightly distorted by the measurements on exit pages, but this is unlikely to have a meaningful impact when examining high level aggregate measurements. Drill downs to specific pages should consider the proportion of exits on that page when evaluating the time on page measurement.

Time on page is often used as a proxy for user engagement, with higher time on page taken to mean higher engagement. This should be carefully considered in context, as high time on page could indicate that a user is having difficulty finding the information they are looking for, and low time on page could also indicate that a user has been able to successfully complete their task quickly.

B.3 Bounce rate

Definition: The proportion of sessions that are considered a 'bounced' session. A 'bounce' is a session where a user views only a single page and does not trigger any events on that page. The bounce rate is calculated as:

$$\frac{(Total\ bounced\ sessions)}{(Total\ sessions)}$$

Constraints: Bounce identification and bounce rate are highly dependent on event configuration, which varies across sites. Bounce rate will be reduced on sites where scroll depth or timing events have been configured.

Considerations: Bounce rate is often used as a proxy for user disengagement, but in some contexts may mean that a user was able to find the information they needed immediately, without needing to further interact with the site.

B.4 Users

Definition: Distinct count of the number of users who visited the site.

Constraints: User identification is determined by a unique identifier contained in a tracking cookie. Individuals who clear cookies, visit the site in 'private browsing mode', visit the site from a different device or browser, or opt-out of google analytics will be counted as multiple users. In circumstances where multiple people are accessing the site from a shared device, they may be identified as a single user.

Considerations: The number of users is frequently treated as a proxy for the number of people accessing the site, but it is important to recognise that one user does not equate to one person.

B.5 New users

Definition: Distinct count of the number of users who visited the site who did not have a tracking cookie indicating that they had previously visited the site. Users who already have a tracking cookie are considered 'returning users'.

Constraints: User status is calculated based on tracking cookies. People who clear cookies, visit the site in 'private browsing mode', visit the site from a different device or browser, or opt-out of google analytics will be counted as new users even if they have previously visited the site. In circumstances where a person is accessing the site from a shared device, they may be counted as a returning user if the device has accessed the site, regardless of whether it is the person's first visit or not.

Considerations: Given the constraint above, it's likely that the number of new users will be overestimated and the number of returning users will be underestimated.

New and returning users are not mutually exclusive – it is possible for a user to be counted as both a new and returning user during a given reporting period.

B.6 Proportion of traffic generated by new users

Definition: The percentage of sessions where the user is categorised a new user. This metric is calculated as:

$$\frac{(Sessions\ with\ new\ user)}{(Total\ sessions)}$$

Constraints: As per new users above.

Considerations: The proportion of traffic generated by new users is likely to be overestimated.

B.7 Traffic by source

Definition: Count of the number of sessions started across different traffic sources. Traffic sources are the way in which a user finds the website. Google analytics has six traffic source categories, detailed in the table below.

Table 7 Traffic source categories

Traffic source	Definition
Direct	The user accesses the website by: • Typing the website URL into the browser address bar • Using a bookmarked link
Organic	The user accesses the website via a service engine (e.g. Google, Bing, or Yahoo)
Referral	The user accesses the website by following a link on another website
Email	The user accesses the website by following a link contained in an email, such as a newsletter
CPM (cost per thousand impressions)	The user accesses the website via a link contained in a marketing campaign (e.g. paid advertisements)
Other	Any access that cannot be categorised as per the above

Considerations: While individual sessions will have only one traffic source, a user may generate multiple sessions and may access the site from a variety of sources.

B.8 Government referrals

Definition: Government referrals are a subset of referral traffic. This is a calculated metric, created by parsing the referral URL to identify referrals from the ".gov.au" domain.

Constraints: Identification of Australian Government or other government referrals relies on the URL string following the convention of including the ".gov.au".

Considerations: A strong theme in the research underpinning 'Difficult to find information' was the issues caused by fragmentation of website content across multiple government websites. Government referral traffic provides a mechanism for identifying content where users are visiting multiple government websites when attempting to reach a goal or complete a task.

B. 9 Proportion of government referrals

Definition: Government referrals are a subset of referral traffic, calculated as:

$$\frac{(Total\ government\ referral\ traffic)}{(Total\ referral\ traffic)}$$

Constraints: As per government referrals above.

Considerations: As per government referrals above.

B.10 Traffic by region

Definition: Count of the number of sessions started, categorised by geographic region. The regions included in this analysis were the eight Australian states and territories.

Constraints: Location data is based on internet protocol (IP) address, and the accuracy can be affected by mobile use, internet service provider configuration, or the

use of virtual private networks, however the impact of any discrepancy is likely to be low.

B.11 Website category

Definition: Website category is an internal categorisation system developed by the Observatory, which classifies subscriber sites as belonging to one of eight sectors:

- Border Control and Immigration
- Business and Employment
- Community Support and Accessibility
- Coronavirus Response
- Defence Force
- Education
- Health
- Science and Technology

Constraints: The classification of the websites was a manual process, and each website was assigned a single category. In some cases, a website may have been related to multiple categories and a primary category was selected. If the Observatory does not have a subscriber from a particular policy area, the area has not been included.

Appendix C: Benchmarks

C.1 Calculation

The benchmarks were calculated using the 'by domain' aggregation table extracted from Google BigQuery. This dataset included domains for all agencies with active Observatory subscriptions between January 1 to March 31 2021 (inclusive).

Prior to computing the benchmark statistics, websites with less than 100 visitors over the three-month period were removed from the dataset. It is likely that sites with such low traffic are non-production websites not yet being used by the public.

Outliers were removed from each distribution before calculating the benchmark statistics for that metric. Outliers are data points that lie considerably beyond the bounds of most of the data in the distribution. They may exist for a number of reasons, such as exceptional traffic or configuration errors. The outliers were removed to make the benchmarks more reflective of the majority of websites.

The method chosen to remove the outliers was to use distance from bounding percentiles based on the interquartile range (IQR). Datapoints were removed where they were 1.5 times the IQR below the 25% percentile or 1.5 times the IQR above the 75% percentile (see formula). This method was chosen because there was a relatively high standard deviation for some of the datasets so standard deviation based measures made the data distribution less normal.

$$Data\ removed = (x < Q(25\%) - IQR \times 1.5)\ OR\ (x > Q(75\%) + IQR \times 1.5)$$
$$where\ IQR = Q(75\%) - Q(25\%)$$

It is important to note that the benchmark averages will differ to the averages referenced in the results of the exploratory research as they are calculated based on different aggregation tables.

C.2 Summary table

Metric	Average	Median	Standard deviation	25th quantile	75th quantile
Average pages per session	7.44	6.59	3.67	4.87	9.05
Average time on page (seconds)	217.60	204.58	107.36	146.03	272.77
Bounce rate (%)	25.91	25.14	13.16	17.44	34.07
Proportion of traffic generated by new users (%)	63.47	64.21	14.53	53.00	72.85
Proportion of government referrals (%)	53.37	52.10	27.44	33.92	75.73