

A guide to auditing your website or intranet

Prepared by Lithos Partners
November 2023



Background notes and introduction

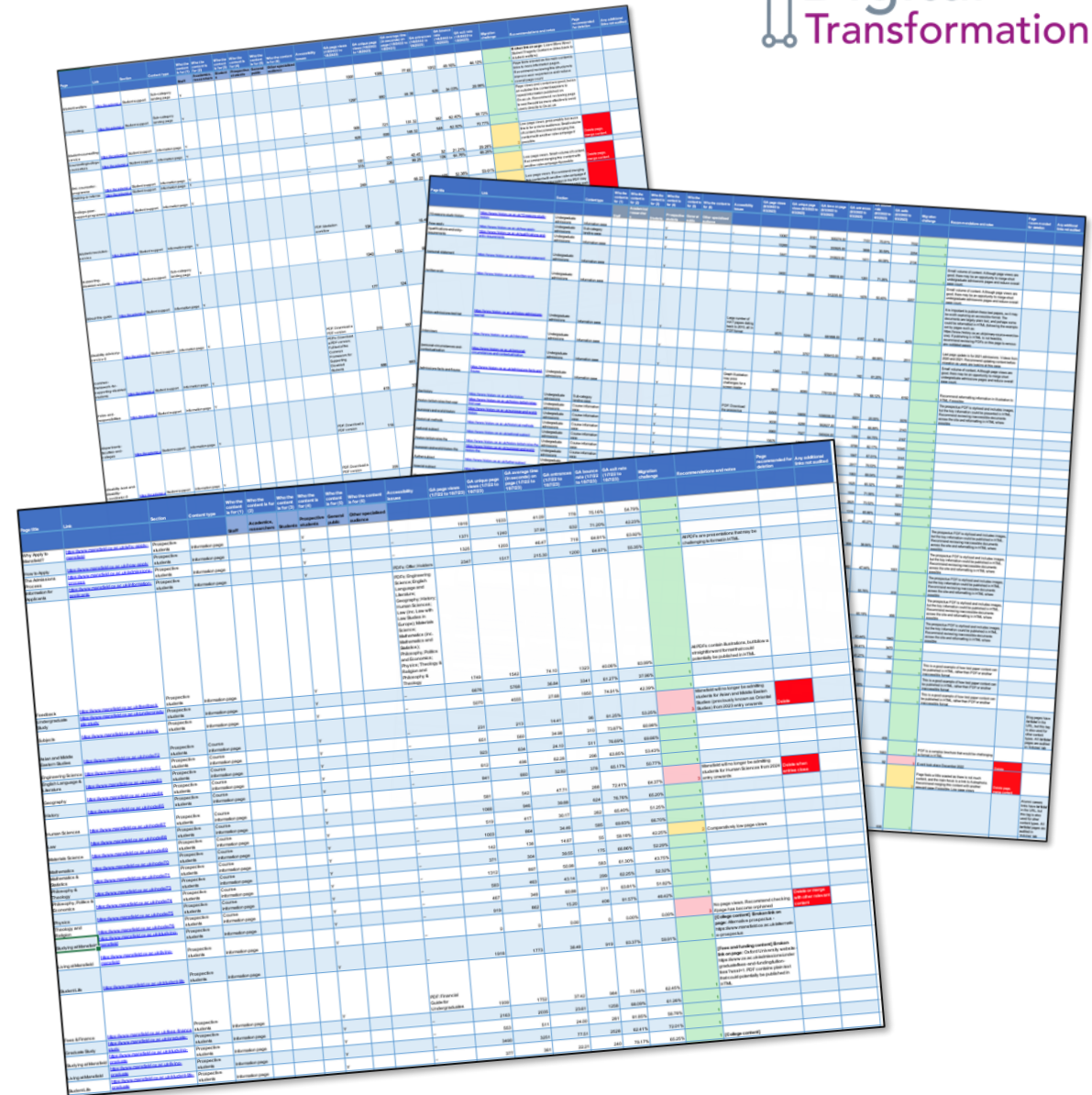
Oxford content audits

Over the long vacation, we audited a sample of Oxford websites and intranets.

The aims of this audit were:

- to understand what content these contained, to inform the content strategy,
- to give us estimate and plan the Mosaic content migration

Doing this also enabled us to develop a repeatable process and templates which any site owner can use to audit their own site.



The image displays three overlapping spreadsheets representing content audit data. Each spreadsheet has a complex header with columns for various metrics and categories. The data rows include website URLs, content types (e.g., 'Page', 'Image', 'Video'), and various numerical values. Some cells are highlighted in green, yellow, or red, indicating different levels of content quality or migration status. The spreadsheets are tilted and layered, showing different sections of the audit data.

Why audit your content?

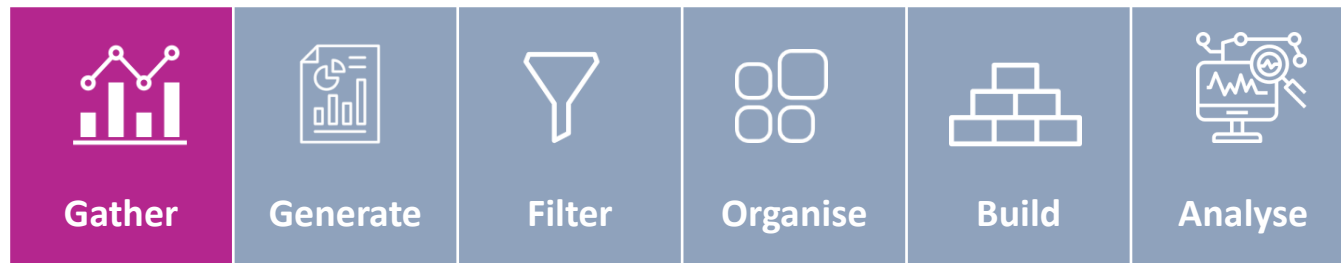
Doing an audit will give you the information you need to

- gauge the overall migration challenge – understand how much content do you have and what work is needed to move it
- identify blocks of pages that need attention before they're moved
- estimate volume of pages on your site that need to be considered during migration
- help you prioritise any future content improvement work

Building an Oxford content audit



Gather data for your content audit



Gathering your data

You'll need a list of pages on your site to get started and there are a few places to look

1

From your Content Management System

Most CMSs at Oxford will allow you to output a list of pages directly

2

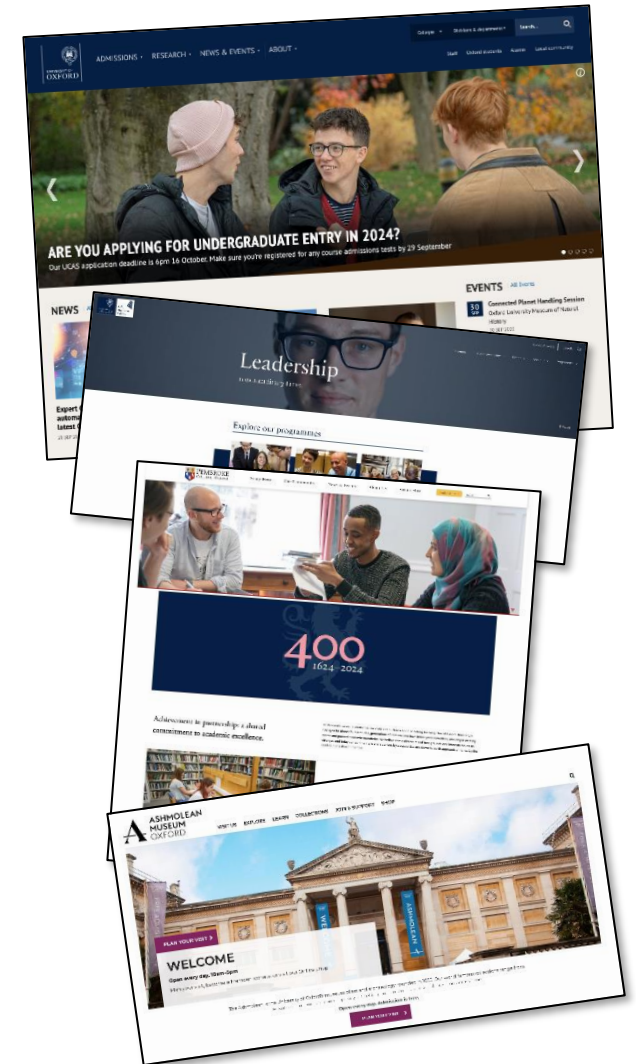
From a site map

A site map will give you a basic list of pages to work from

3

From Google Analytics

If you have an analytics tool, such as Google Analytics (GA4), this should also give you a list of all the pages it is tracking



Gathering your data

If you're on a Mosaic site, you can use the following links to generate a list of pages:

[site_prefix].web.ox.ac.uk/sitelist.html

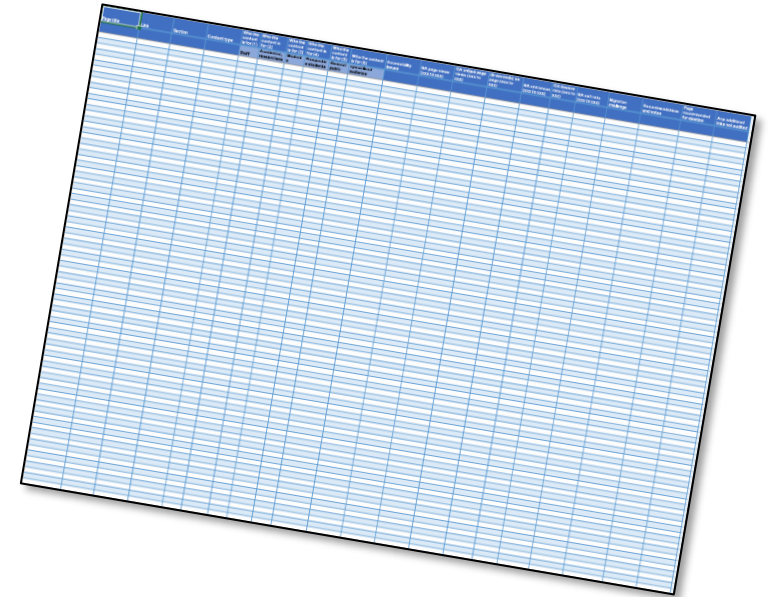
[site_prefix].web.ox.ac.uk/sitemap.xml

[custom_domain]/sitemap.xml or [custom_domain]/sitelist.html

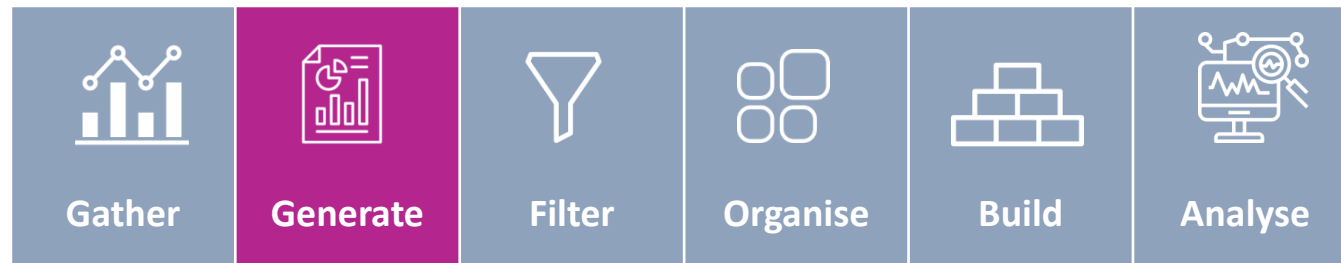
Your content audit template

To begin the content auditing process, you'll have a few documents to help you:

- blank template to record your content audit
- completed content audit from another Oxford site, which you can refer to when working on your audit
- This step-by-step guide to completing your audit



Generate a Google Analytics (GA4) report

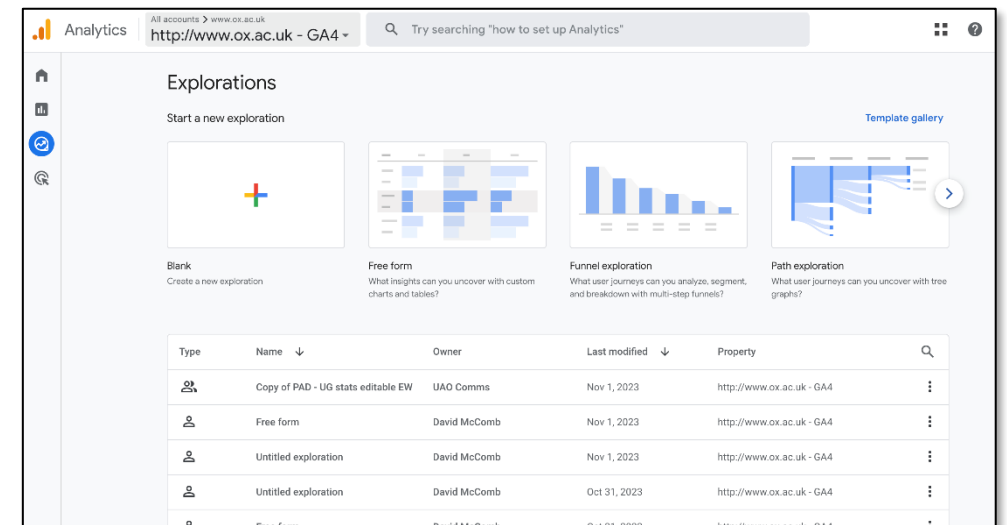


Using Google Analytics 4 (GA4)

If you do not have access to a GA4 dashboard for your website, you should request this as soon as possible. If you use another analytics package, please refer to their online guides for assistance.

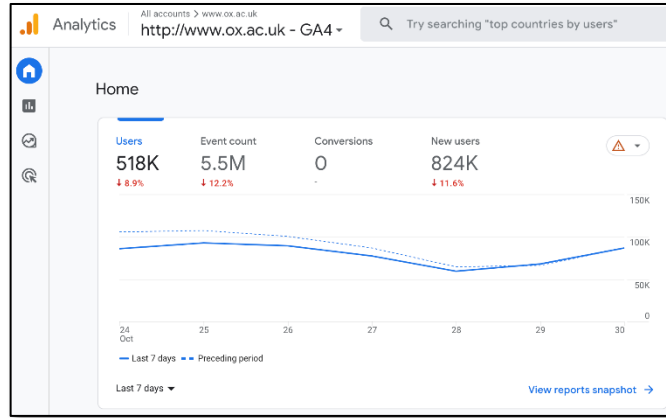
If you have used Google Analytics in the past, but not GA4, the dashboard may be confusing. But this guide will show you how to create a report for your content audit.

Google Analytics is capable of amazing things and can offer valuable user insights. But you only need a basic GA4 report to begin understanding the volume of published pages and the migration challenge.

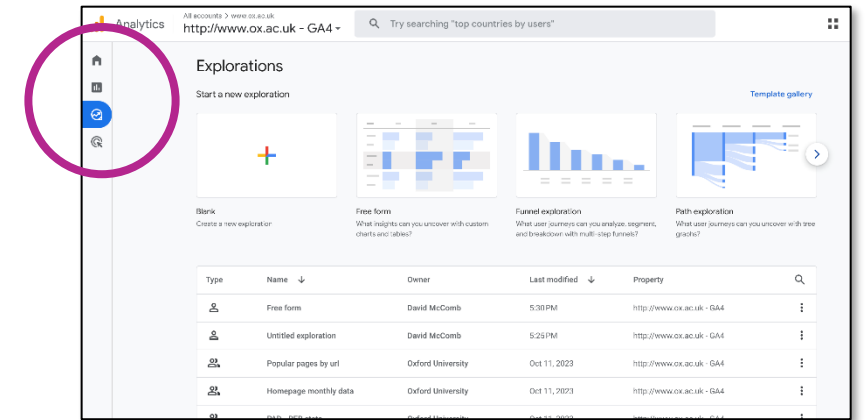


Generating a custom GA4 report

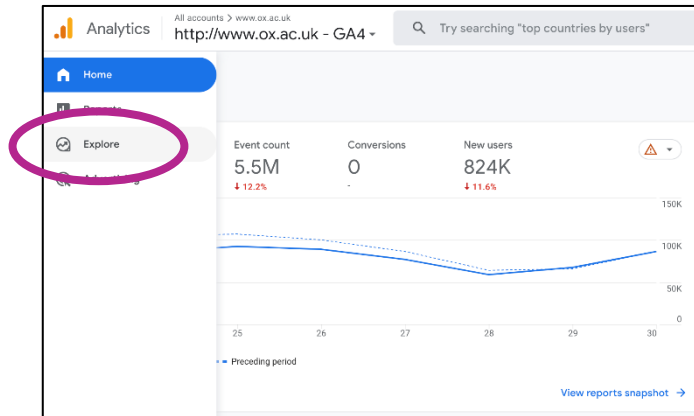
1 Log in to your website's GA4 account



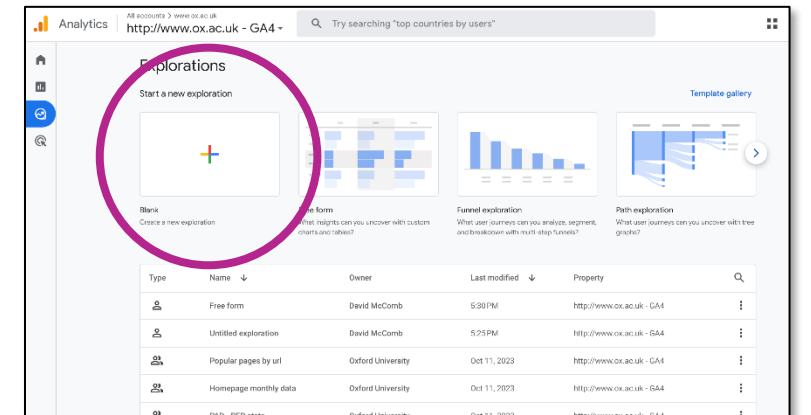
2 Go top, left-hand corner of the screen and roll over the icons



3 Select 'Explore'



4 Choose a blank report on the following menu screen



Setting up your custom GA4 report

1

Exploration name

Add a name for your report

2

Technique

Check this is set to 'Free form'

3

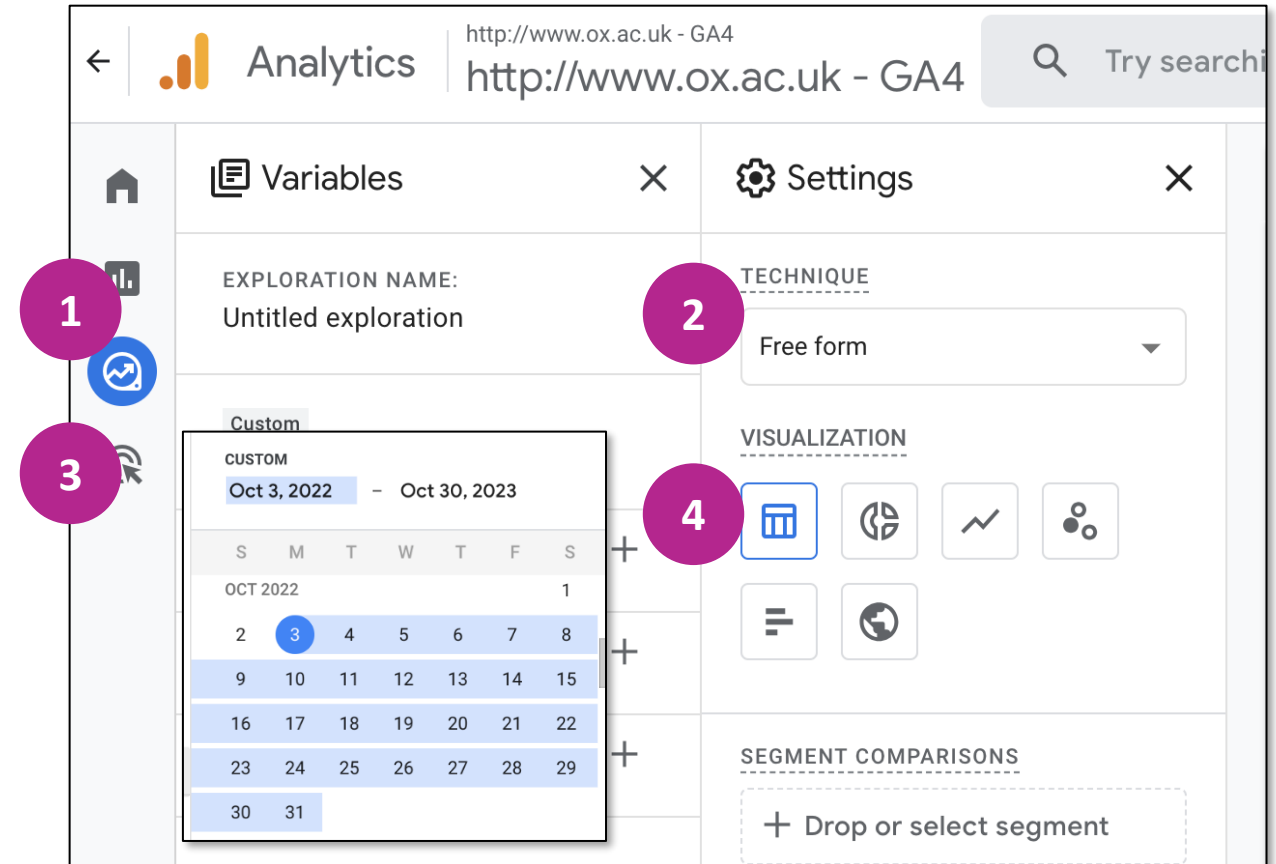
Date range: Custom

To collect comprehensive data to inform and support content migration decisions, GA4 reports should cover a 12-month period

4

Visualization

Select the table icon



Analytics | http://www.ox.ac.uk - GA4

Variables × Settings ×

EXPLORATION NAME: Untitled exploration

TECHNIQUE: Free form

Custom

CUSTOM

Oct 3, 2022 - Oct 30, 2023

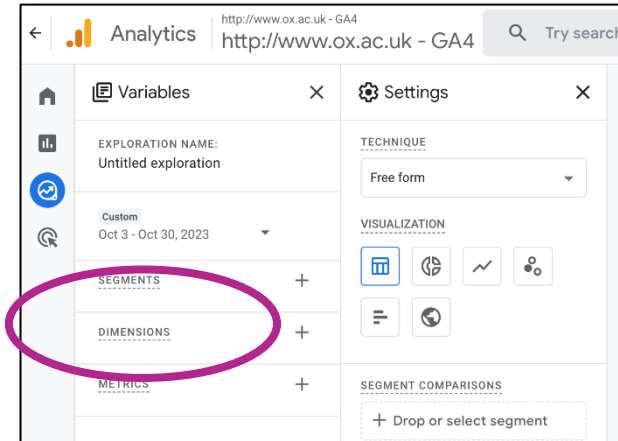
S	M	T	W	T	F	S
						1
2	3	4	5	6	7	8
9	10	11	12	13	14	15
16	17	18	19	20	21	22
23	24	25	26	27	28	29
30	31					

VISUALIZATION: [Table icon] [Pie chart icon] [Line chart icon] [Map icon]

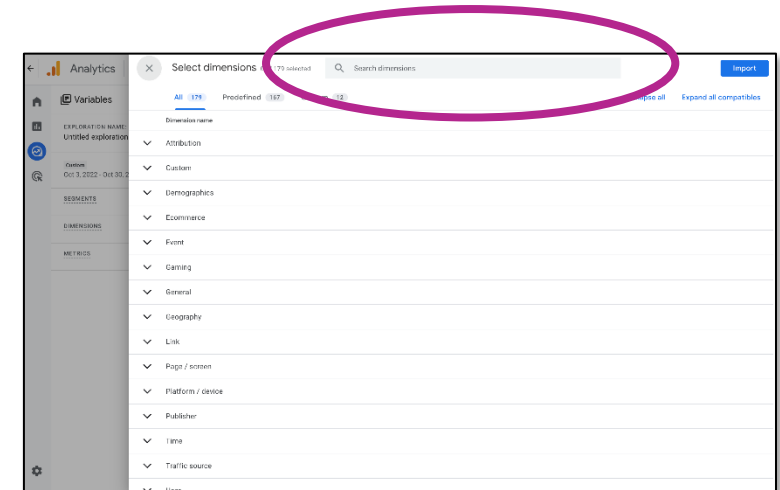
SEGMENT COMPARISONS: + Drop or select segment

Selecting data for audit rows

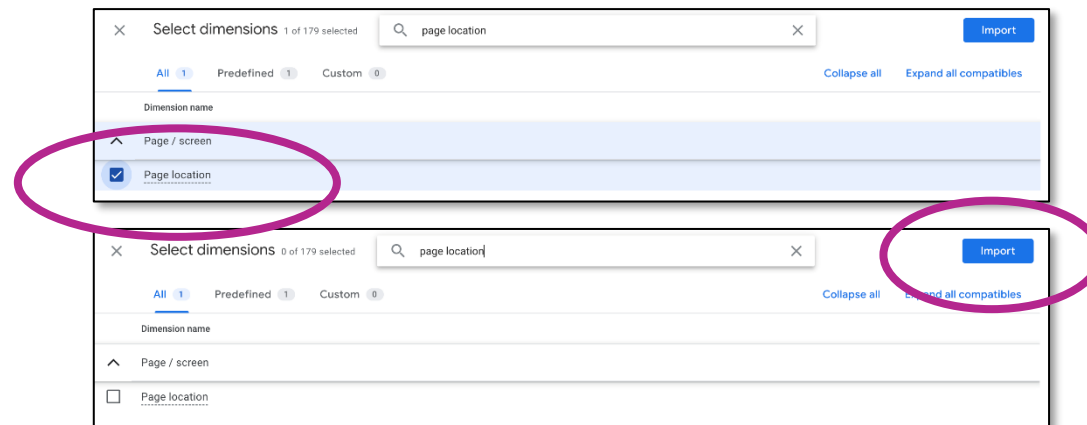
1 Select the 'Dimensions' menu



2 Search for 'Page location' in the search bar at the top of the page

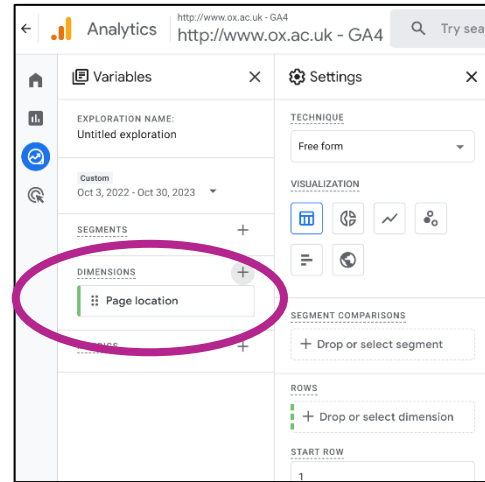


3 When 'Page location' appears in the list below the search bar, tick the box and select 'Import'

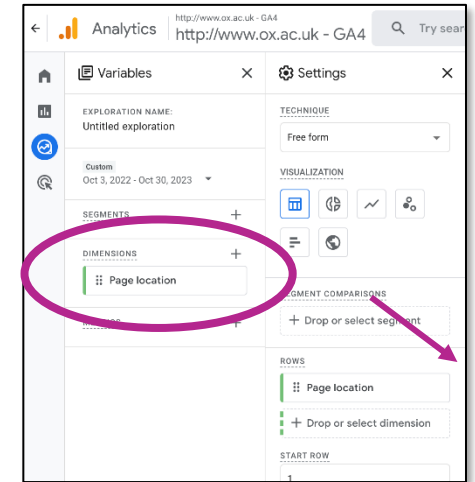


Selecting data for audit rows

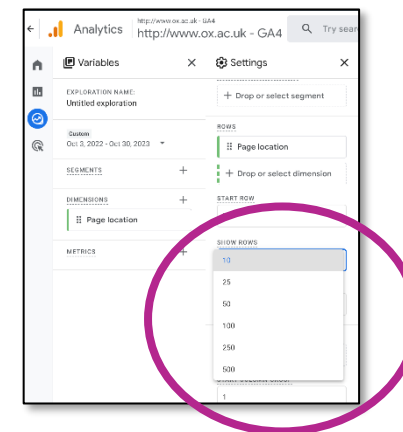
4 You will see that a 'Page location' box has appeared under 'Dimensions'



5 Click on the 'Page location' box and drag it to 'Rows' in the 'Settings' column. The 'Page location' box should appear under 'Rows'

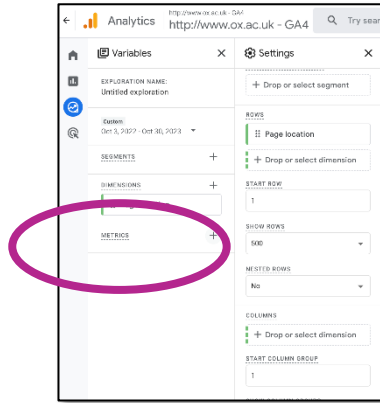


6 At this stage you should also select 'Show rows' and choose '500'. This will display 500 lines of data in your GA4 online report – but the report you download will have many more lines of data to analyse

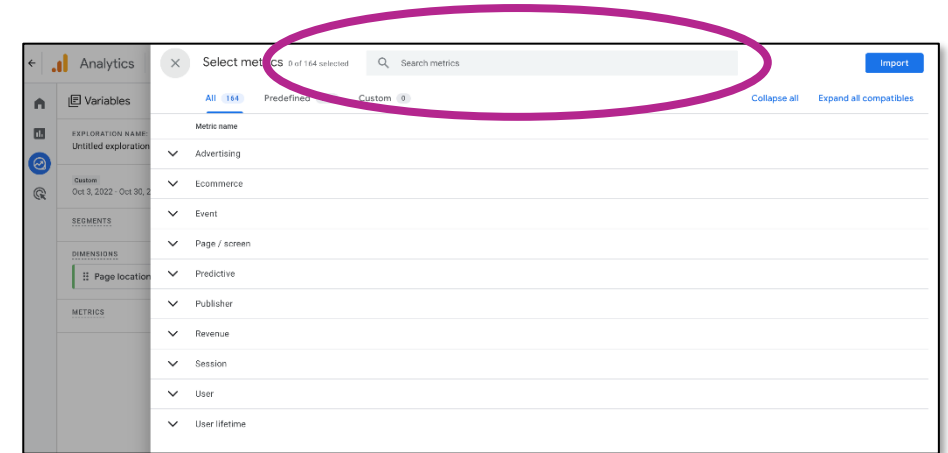


Selecting metrics for your GA4 report

1 Select the 'Metrics' menu



2 As you did for page location, search for and add metrics



3 Select and add these metrics:

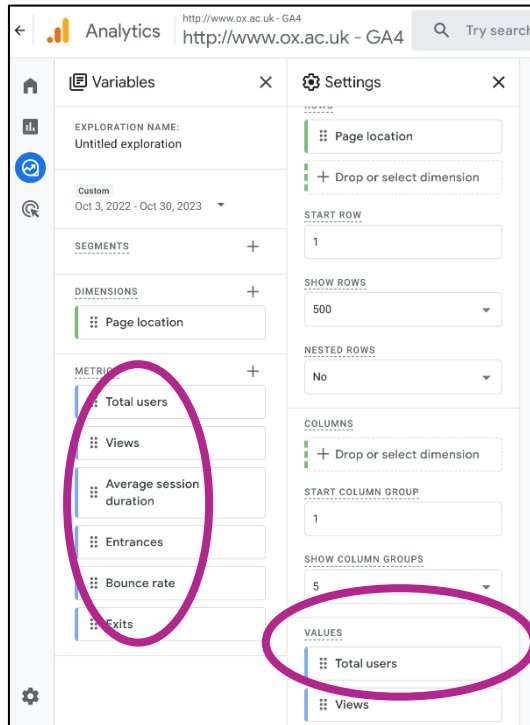
- total users
- views
- average session duration
- entrances
- bounce rate
- exits

4 When the metric you need appears in the menu, check the tick box and hit 'Import'



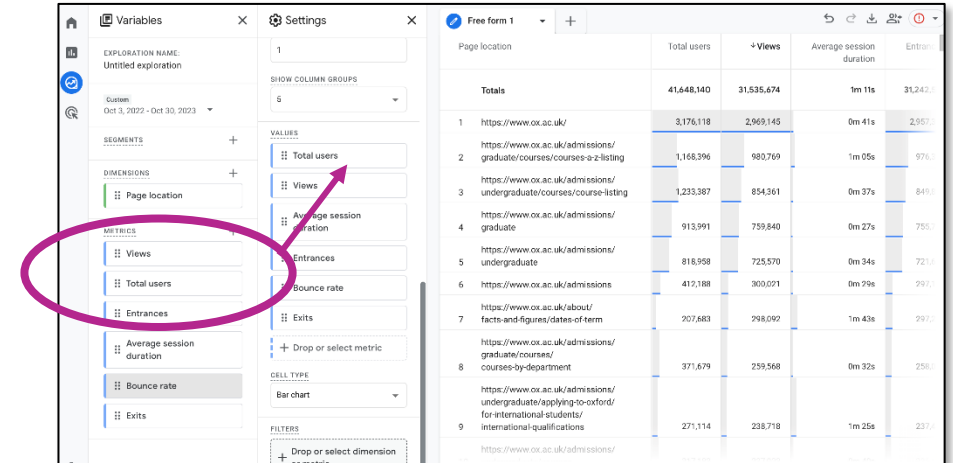
Adding metrics to your GA4 report

1 You will see that boxes for each of your metrics have appeared under the 'Metrics' header



Scroll down the 'Settings' column until you see a section with the header 'Values'

2 Drag the metrics boxes to 'Values' and you'll see your audit starting to build

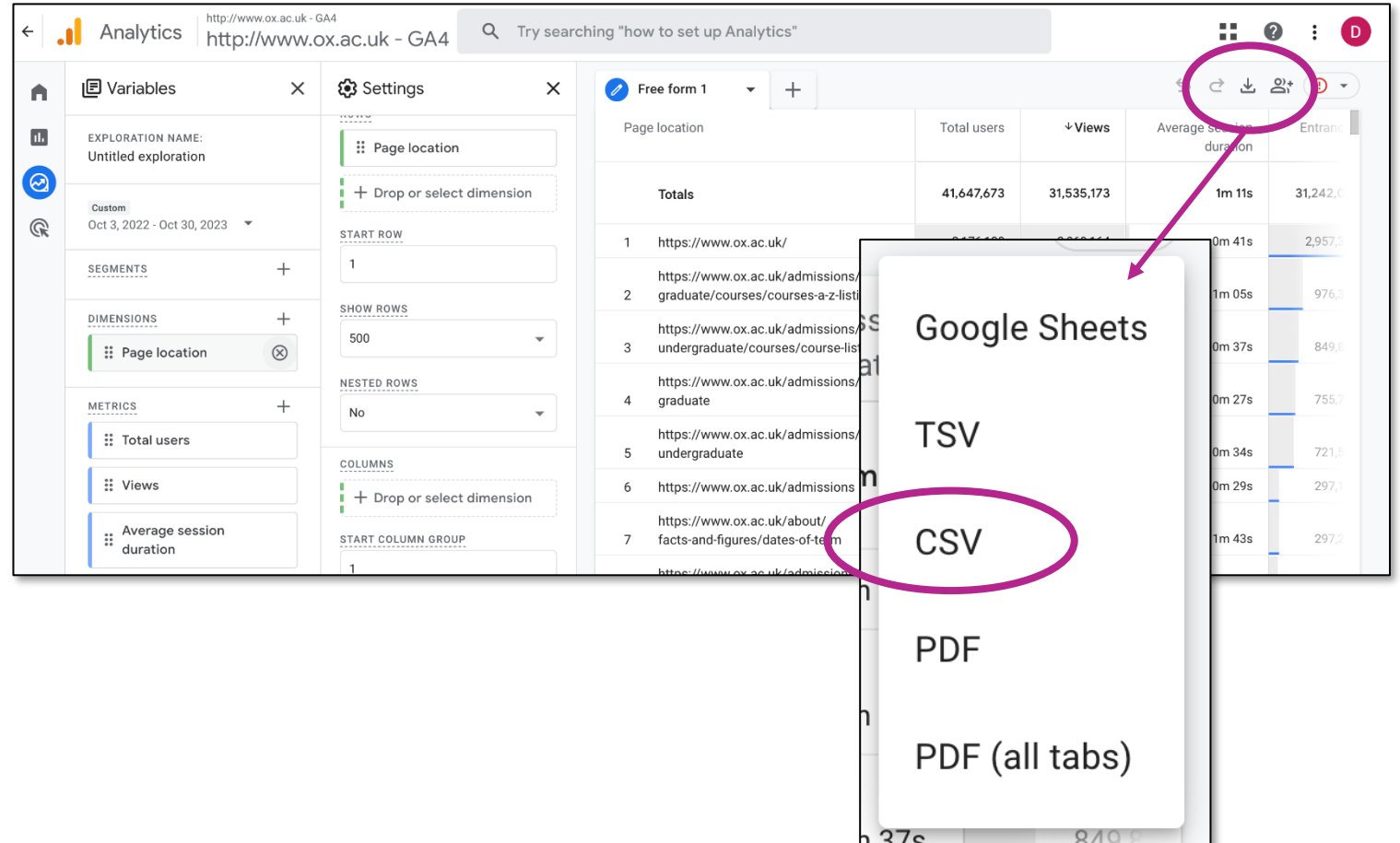


In the audit you download, the metrics will appear in the same order (from left to right) as you place these boxes in the 'Values' column (that is, the first metric in the spreadsheet will be the top box in 'Values', the second metric will be the second box and so on). You can move the metric boxes in GA4 to change the order

Downloading your GA4 report

To download your report, select the download icon at the top, right-hand corner of the GA4 dashboard.

Choose 'CSV' to download data that can be easily inserted into your content audit template.



The screenshot shows the Google Analytics GA4 interface. On the right, a table displays report data with columns for Page location, Total users, Views, Average session duration, and Entrances. A download icon (a downward arrow) in the top right corner is circled in pink. A dropdown menu is open, listing download options: Google Sheets, TSV, CSV (circled in pink), PDF, and PDF (all tabs). An arrow points from the download icon to the menu.

Page location	Total users	Views	Average session duration	Entrances
Totals	41,647,673	31,535,173	1m 11s	31,242,000
1 https://www.ox.ac.uk/	2,175,122	2,222,111	0m 41s	2,957,000
2 https://www.ox.ac.uk/admissions/graduate/courses/courses-a-z-list	1,111,111	1,111,111	1m 05s	976,000
3 https://www.ox.ac.uk/admissions/undergraduate/courses/course-list	1,111,111	1,111,111	0m 37s	849,000
4 https://www.ox.ac.uk/admissions/graduate	1,111,111	1,111,111	0m 27s	755,000
5 https://www.ox.ac.uk/admissions/undergraduate	1,111,111	1,111,111	0m 34s	721,000
6 https://www.ox.ac.uk/admissions	1,111,111	1,111,111	0m 29s	297,000
7 https://www.ox.ac.uk/about/facts-and-figures/dates-of-the-university	1,111,111	1,111,111	1m 43s	297,000

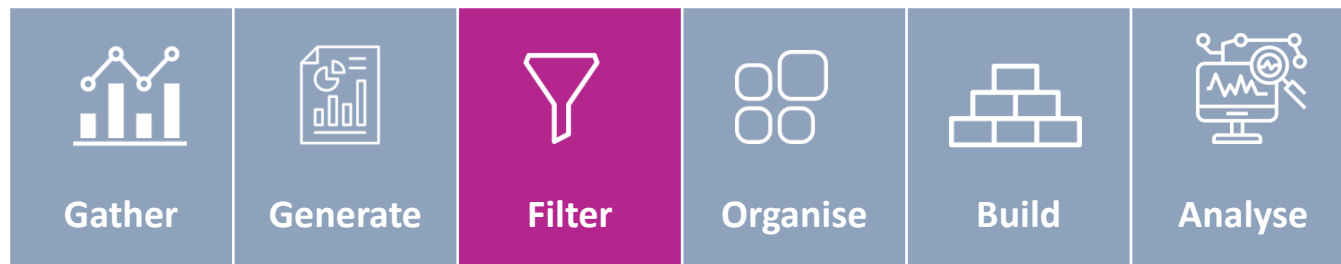
Insert raw data into your audit template

In case you need to refer to the GA4 raw data during analysis, paste this into the GA4 data (raw) tab in the audit template.

Also keep a separate copy of the GA4 report so that it can be filtered

Page	GA4 total users (1/6/22 to 14/7/23)	GA4 views (1/6/22 to 14/7/23)	GA4 average session duration (seconds) (1/6/22 to 14/7/23)	GA4 entrances (1/6/22 to 14/7/23)	GA4 % bounce rate (1/6/22 to 14/7/23)	GA4 exits (1/6/22 to 14/7/23)
https://ox.ac.uk/	2914301	3036618	41.04	3023707	62.42%	3E+06
https://oxford.shorthandstories.com/BrainWaves-in-search-of-better-teenage-mental-health/index.html	465	646	50.13899302	516	0.4280443	516
https://oxford.shorthandstories.com/a-history-of-neuroscience-at-oxford/index.html	1704	1833	148.3089347	1704	0.499731	1627
https://oxford.shorthandstories.com/brain-alzheimers/	77	77	0	77	1	77
https://oxford.shorthandstories.com/brain-alzheimers/index.html	1472	1601	83.01988398	1446	0.6333118	1420
https://oxford.shorthandstories.com/brain-and-mental-health-from-infancy-to-adolescence/	749	981	86.41894254	878	0.6921549	904
https://oxford.shorthandstories.com/brain-and-mental-health-innovation-at-oxford/index.html	1239	1782	204.0355168	1420	0.6460072	1498
https://oxford.shorthandstories.com/brain-andy-przybylski-uncovering-the-impact-of-technology-on-mental-health/index.html	207	258	372.5066137	232	0.4542254	232
https://oxford.shorthandstories.com/brain-cathy-creswell-supporting-the-mental-health-of-families/index.html	52	52	37.55985117	52	0.5	52
https://oxford.shorthandstories.com/brain-language-literature/index.html	516	620	165.4933061	516	0.5988372	491
https://oxford.shorthandstories.com/brain-laurence-wainwright-recognising-the-impact-of-the-climate-on-mental-health/	52	52	181.9291181	52	0.3246753	26
https://oxford.shorthandstories.com/cancer-prevention/	181	465	246.6334625	336	0.4281768	284
https://oxford.shorthandstories.com/cancer-prevention/?utm_source=linkedin&utm_medium=referral&utm_campaign=cancer	52	207	225.2459914	181	0.8563536	181
https://oxford.shorthandstories.com/cancer-prevention/?utm_source=twitter&utm_medium=referral&utm_campaign=cancer	77	77	4.918536571	77	0.6623377	77
https://oxford.shorthandstories.com/dementia/index.html	2117	14692	1525.593469	3744	0.5752974	3718
https://oxford.shorthandstories.com/disease-and-degeneration/index.html	2014	2505	104.3341189	2221	0.6489493	2117
https://oxford.shorthandstories.com/epilepsy/index.html	826	955	185.6086184	775	0.6173121	775
https://oxford.shorthandstories.com/gender-and-sexual-identity/index.html	2453	2815	150.0123252	2556	0.4860659	2505
https://oxford.shorthandstories.com/innovation-dice/index.html	800	1291	131.9839491	1188	0.5738881	1188
https://oxford.shorthandstories.com/innovation-first-light-fusion/index.html	362	594	331.827765	362	0.6313646	362
https://oxford.shorthandstories.com/innovation-rollsroyce/index.html	52	77	6.88046589	77	0.6623377	77
https://oxford.shorthandstories.com/innovation-vaccine/index.html	258	258	77.98146937	258	0.1007752	258
https://oxford.shorthandstories.com/innovation-virtual-therapy-made-real/index.html	387	568	281.6327231	387	0.5784114	387
https://oxford.shorthandstories.com/maths-plus-cancer/	5216	8340	96.44022698	7540	0.6429345	7540
https://oxford.shorthandstories.com/maths-plus-cancer/?fbclid=PAaAaYurshcEVNk4yeRk6R0EPFNZ6XY7ovPwiVsrFbfaQjPZo6_EFs24qr-oL0	52	52	0	52	1	52
https://oxford.shorthandstories.com/maths-plus-cancer/index.html?utm_source=linkedin&utm_medium=referral&utm_campaign=cancer	129	155	18.72638979	155	0.7127072	155
https://oxford.shorthandstories.com/maths-plus-cancer/index.html?utm_source=twitter&utm_medium=referral&utm_campaign=cancer	516	646	5.412371633	568	0.8703704	568
https://oxford.shorthandstories.com/mental-health-and-the-world-around-us/	3176	3899	109.316835	3357	0.6362405	3408
https://oxford.shorthandstories.com/mental-health-data/index.html	362	516	53.15414937	387	0.790224	413
https://oxford.shorthandstories.com/parental-and-maternal/	1239	1601	159.5551228	1239	0.6933167	1239
https://oxford.shorthandstories.com/professor-kj-patel/	3047	3899	147.8832023	3615	0.6644562	3615
https://oxford.shorthandstories.com/professor-kj-patel/index.html	129	232	244.1500896	155	0.6280193	207

Filter your GA4 report



Removing redundant rows

Depending on the size of your site, you will now have an Excel report that may contain thousands of rows of data

But it's likely a large portion of these rows are not relevant to the content auditing exercise.

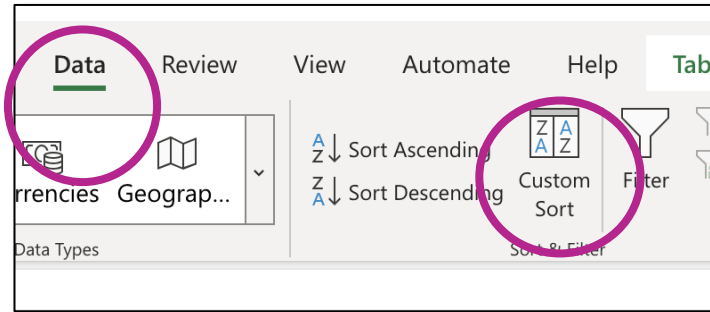
Before starting to build your live audit, remove these redundant rows of data to make your content audit easier to manage.

Page	GA4 total users (1/6/22 to 14/7/23)	GA4 views (1/6/22 to 14/7/23)	GA4 average session duration (seconds) (1/6/22 to 14/7/23)	GA4 entrances (1/6/22 to 14/7/23)	GA4 % bounce rate (1/6/22 to 14/7/23)	GA4 exits (1/6/22 to 14/7/23)
https://ox.ac.uk/	2914301	3036618	41.04	3023707	62.42%	3E+06
https://oxford.shorthandstories.com/BrainWaves-in-search-of-better-teenage-mental-health/index.html	465	646	50.13899302	516	0.4280443	516
https://oxford.shorthandstories.com/a-history-of-neuroscience-at-oxford/index.html	1704	1833	148.3089347	1704	0.499731	1627
https://oxford.shorthandstories.com/brain-alzheimers/	77	77	0	77	1	77
https://oxford.shorthandstories.com/brain-alzheimers/index.html	1472	1601	83.01988398	1446	0.6333118	1420
https://oxford.shorthandstories.com/brain-and-mental-health-from-infancy-to-adolescence/	749	981	86.41894254	878	0.6921549	904
https://oxford.shorthandstories.com/brain-and-mental-health-innovation-at-oxford/index.html	1239	1782	204.0355168	1420	0.6460072	1498
https://oxford.shorthandstories.com/brain-andy-przybylski-uncovering-the-impact-of-technology-on-mental-health/	207	258	372.5066137	232	0.4542254	232
https://oxford.shorthandstories.com/brain-cathy-creswell-supporting-the-mental-health-of-families/index.html	52	52	37.55985117	52	0.5	52
https://oxford.shorthandstories.com/brain-language-literature/index.html	516	620	165.4933061	516	0.5988372	491
https://oxford.shorthandstories.com/brain-laurence-wainwright-recognising-the-impact-of-the-climate-on-mental-health/	52	52	181.9291181	52	0.3246753	26
https://oxford.shorthandstories.com/cancer-prevention/	181	465	246.6334625	336	0.4281768	284
https://oxford.shorthandstories.com/cancer-prevention/?utm_source=linkedin&utm_medium=referral&utm_campaign=...	52	207	225.2459914	181	0.8563536	181
https://oxford.shorthandstories.com/cancer-prevention/?utm_source=twitter&utm_medium=referral&utm_campaign=...	77	77	4.918536571	77	0.6623377	77
https://oxford.shorthandstories.com/dementia/index.html	2117	14692	1525.593469	3744	0.5752974	3718
https://oxford.shorthandstories.com/disease-and-degeneration/index.html	2014	2505	104.3341189	2221	0.6489493	2117
https://oxford.shorthandstories.com/epilepsy/index.html	826	955	185.6086184	775	0.6173121	775
https://oxford.shorthandstories.com/gender-and-sexual-identity/index.html	2453	2815	150.0123252	2556	0.4860659	2505
https://oxford.shorthandstories.com/innovation-dice/index.html	800	1291	131.9839491	1188	0.5738881	1188
https://oxford.shorthandstories.com/innovation-first-light-fusion/index.html	362	594	331.827765	362	0.6313646	362
https://oxford.shorthandstories.com/innovation-roll-royce/index.html	52	77	6.88046589	77	0.6623377	77
https://oxford.shorthandstories.com/innovation-vaccine/index.html	258	258	77.98146937	258	0.1007752	258
https://oxford.shorthandstories.com/innovation-virtual-therapy-made-real/index.html	387	568	281.6327231	387	0.5784114	387
https://oxford.shorthandstories.com/math-plus-cancer/	5216	8340	96.44022698	7540	0.6429345	7540
https://oxford.shorthandstories.com/math-plus-cancer/?fbclid=PAAaYurshcEVNK4yeRk6R0EPFNZ6XY7ovPwiVS...	52	52	0	52	1	52
https://oxford.shorthandstories.com/math-plus-cancer/index.html?utm_source=linkedin&utm_medium=referral&utm_campaign=...	129	155	18.72638979	155	0.7127072	155
https://oxford.shorthandstories.com/math-plus-cancer/index.html?utm_source=twitter&utm_medium=referral&utm_campaign=...	516	646	5.412371633	568	0.8703704	568
https://oxford.shorthandstories.com/mental-health-and-the-world-around-us/	3176	3899	109.316835	3357	0.6362405	3408
https://oxford.shorthandstories.com/mental-health-data/index.html	362	516	53.15414937	387	0.790224	413
https://oxford.shorthandstories.com/parental-and-maternal/	1239	1601	159.5551228	1239	0.6933167	1239
https://oxford.shorthandstories.com/professor-kj-patel/	3047	3899	147.8832023	3615	0.6644562	3615
https://oxford.shorthandstories.com/professor-kj-patel/index.html	129	232	244.1500896	155	0.6280193	207

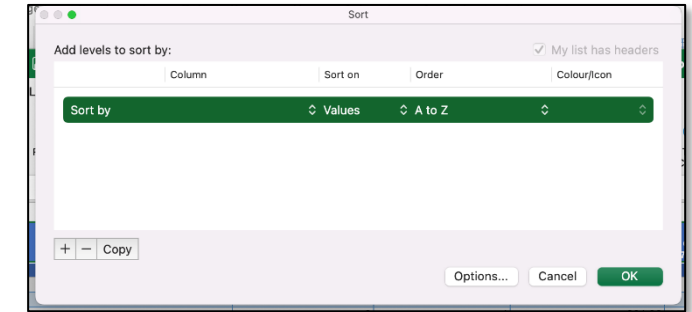
Order data by views

First view results by 'total users'. This will help to push groups of redundant results to the bottom of the sheet, making them easier to delete in groups

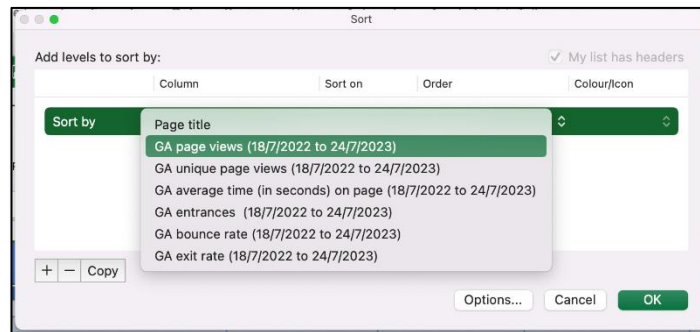
1 In the toolbar select the 'Data' tab and then choose 'Custom Sort'



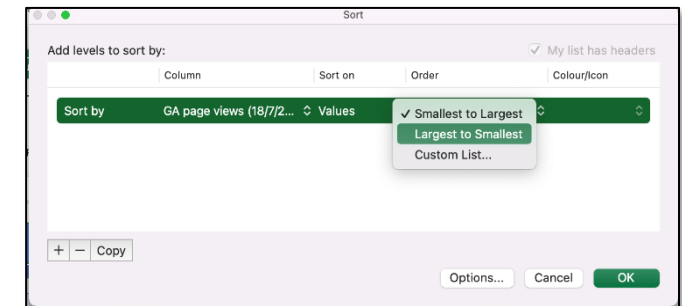
2 The menu allows you to select which column to order



3 Select the 'total users' column



4 View by 'Largest to smallest' and hit OK



Identifying and deleting redundant rows

As the goal of the audit is to record live content pages and understand any migration challenges, there are GA4 results that are not relevant to this exercise, including:

- search results
- listing pages
- translated content
- any URLs that contain unusual characters

Identifying and deleting redundant rows

Examples of what redundant lines look like in a GA report:

<https://oxford.shorthandstories.com/maths-plus-cancer/?fbclid=PAAaYurshcEVNK4yeRk6R0EPFNZ6XY7>

In this case, the URL <https://oxford.shorthandstories.com/maths-plus-cancer> would be recorded in the audit, but this one with the referral string should not.

Rows in the GA4 report that repeat the URL with unusual characters are likely to be erroneous or repeat data and should be deleted. For example:

https://www-ox-ac-uk.translate.google/?_x_tr_sl=en&_x_tr_tl=ar&_x_tr_hl=ar&_x_tr_pto=sc

This is a link to a translated version of the page

<https://www.ox.ac.uk/?from=edurank.org>

This shows a referral to the page

<https://www.ox.ac.uk/about/facts-and-figures/admissions-statistics%20%20%20%20%20>

This is recording additional characters after the URL

<https://www.ox.ac.uk/Events-list?page=2>

This is the second of multiple pages

In the final example, the main <https://www.ox.ac.uk/Events-list> listings page can be recorded, but any subsequent pages (page=2, page=3) can be deleted.

Identifying and deleting redundant rows

Identifying redundant lines becomes easier as you work through the report. Common sense will guide you when spotting lines that are not relevant in the audit

Common characters in redundant rows

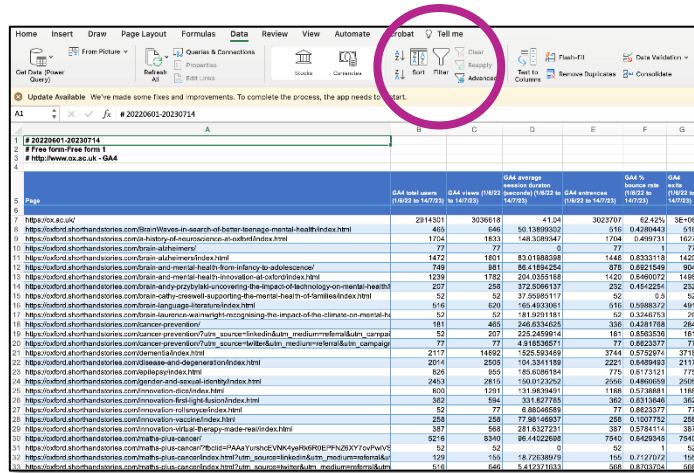
Equals:	=
Underscore:	_
Ampersand:	&
Question mark:	?

You may need to repeat this filtering step a few times and use different characters to highlight and delete all redundant rows in your GA4 report. Even after doing this, a handful of redundant rows may need to be removed when building the audit.

Identifying and deleting redundant rows

1

Click on the first cell of the column you want to filter, then select the 'Filter' funnel icon in the Data tab



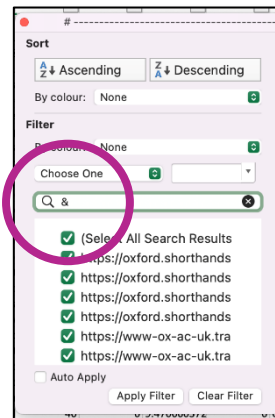
2

Drop-down arrows will appear at the top of all columns. Click the arrow for the column you want to filter

	GA4 total users (1/6/22 to 14/7/23)	GA4 views (1/6/22 to 14/7/23)	GA4 average session duration (seconds) (1/6/22 to 14/7/23)	GA4 entrances (1/6/22 to 14/7/23)	GA4 % bounce rate (1/6/22 to 14/7/23)	GA4 exits (1/6/22 to 14/7/23)
	2914301	3036618	41.04	3023707	62.42%	3E+06
	465	646	50.13899302	516	0.4280443	516
	1704	1833	148.3089347	1704	0.499731	1627
	77	77	0	77	1	77
	1472	1601	83.01988398	1446	0.6333118	1420
	749	981	86.41894254	878	0.6921549	904
	1239	1782	204.0355168	1420	0.6460072	1498
al-health/i	207	258	372.5066137	232	0.4542254	232
tml	52	52	37.55985117	52	0.5	52
	516	620	165.4933061	516	0.5988372	491
mental-he	52	52	181.9291181	52	0.3246753	26
	181	465	246.6334625	336	0.4281768	284
campai	52	207	225.2459914	181	0.8563536	181

3

In the menu box, choose 'select all' and enter your search term (for example, = or &)



4

Apply the filter to reveal redundant lines

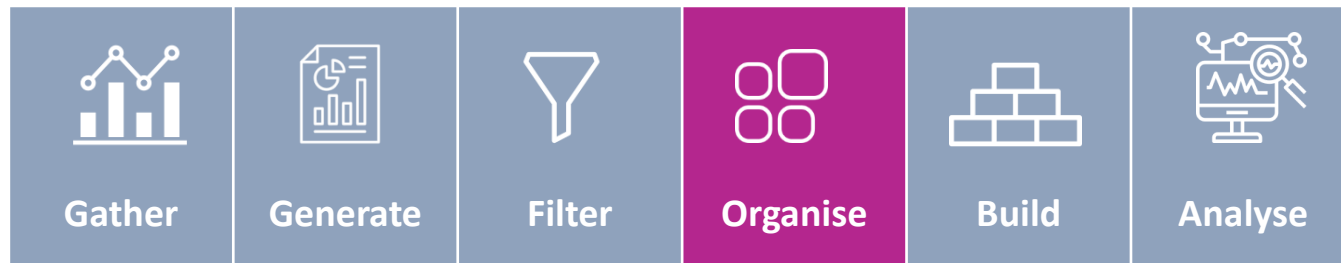
https://www-ox-ac-uk.translate.google/?_x_tr_sl=en&_x_tr_tl=ar&_x_tr_hl=ar&_x_tr_pto=sc
https://www-ox-ac-uk.translate.google/?_x_tr_sl=en&_x_tr_tl=de&_x_tr_hl=de&_x_tr_pto=sc
https://www-ox-ac-uk.translate.google/?_x_tr_sl=en&_x_tr_tl=es&_x_tr_hl=es&_x_tr_pto=sc
https://www-ox-ac-uk.translate.google/?_x_tr_sl=en&_x_tr_tl=es&_x_tr_hl=es-419&_x_tr_pto=sc
https://www-ox-ac-uk.translate.google/?_x_tr_sl=en&_x_tr_tl=fa&_x_tr_hl=fa&_x_tr_pto=sc
https://www-ox-ac-uk.translate.google/?_x_tr_sl=en&_x_tr_tl=fr&_x_tr_hl=fr&_x_tr_pto=sc
https://www-ox-ac-uk.translate.google/?_x_tr_sl=en&_x_tr_tl=hi&_x_tr_hl=hi&_x_tr_pto=sc
https://www-ox-ac-uk.translate.google/?_x_tr_sl=en&_x_tr_tl=id&_x_tr_hl=id&_x_tr_pto=sc
https://www-ox-ac-uk.translate.google/?_x_tr_sl=en&_x_tr_tl=ja&_x_tr_hl=ja&_x_tr_pto=op,sc
https://www-ox-ac-uk.translate.google/?_x_tr_sl=en&_x_tr_tl=ja&_x_tr_hl=ja&_x_tr_pto=sc
https://www-ox-ac-uk.translate.google/?_x_tr_sl=en&_x_tr_tl=nl&_x_tr_hl=nl&_x_tr_pto=sc
https://www-ox-ac-uk.translate.google/?_x_tr_sl=en&_x_tr_tl=pl&_x_tr_hl=pl&_x_tr_pto=sc
https://www-ox-ac-uk.translate.google/?_x_tr_sl=en&_x_tr_tl=pt&_x_tr_hl=pt&_x_tr_pto=sc
https://www-ox-ac-uk.translate.google/?_x_tr_sl=en&_x_tr_tl=pt&_x_tr_hl=pt-PT&_x_tr_pto=sc
https://www-ox-ac-uk.translate.google/?_x_tr_sl=en&_x_tr_tl=ru&_x_tr_hl=ru&_x_tr_pto=sc
https://www-ox-ac-uk.translate.google/?_x_tr_sl=en&_x_tr_tl=ru&_x_tr_hl=ru&_x_tr_pto=wapp
https://www-ox-ac-uk.translate.google/?_x_tr_sl=en&_x_tr_tl=tr&_x_tr_hl=tr&_x_tr_pto=sc
https://www-ox-ac-uk.translate.google/?_x_tr_sl=en&_x_tr_tl=zh-CN&_x_tr_hl=zh-CN&_x_tr_pto=sc

Identifying and deleting redundant rows

Once all redundant lines are removed, you will be left with a GA4 report that only contains 'clean' URLs and the GA4 data attached to them

Page	GA4 total users (1/6/22 to 14/7/23)	GA4 views (1/6/22 to 14/7/23)	GA4 average session duration (seconds) (1/6/22 to 14/7/23)	GA4 entrances (1/6/22 to 14/7/23)	GA4 % bounce rate (1/6/22 to 14/7/23)	GA4 exits (1/6/22 to 14/7/23)
https://ox.ac.uk/	2914301	3036618	41.04	3023707	62.42%	3E+06
https://oxford.shorthandstories.com/BrainWaves-in-search-of-better-teenage-mental-health/index.html	465	646	50.13899302	516	0.4280443	516
https://oxford.shorthandstories.com/a-history-of-neuroscience-at-oxford/index.html	1704	1833	148.3089347	1704	0.499731	1627
https://oxford.shorthandstories.com/brain-alzheimers/	77	77	0	77	1	77
https://oxford.shorthandstories.com/brain-alzheimers/index.html	1472	1601	83.01988398	1446	0.6333118	1420
https://oxford.shorthandstories.com/brain-and-mental-health-from-infancy-to-adolescence/	749	981	86.41894254	878	0.6921549	904
https://oxford.shorthandstories.com/brain-and-mental-health-innovation-at-oxford/index.html	1239	1782	204.0355168	1420	0.6460072	1498
https://oxford.shorthandstories.com/brain-andy-przybylski-uncovering-the-impact-of-technology-on-mental-health/	207	258	372.5066137	232	0.4542254	232
https://oxford.shorthandstories.com/brain-cathy-creswell-supporting-the-mental-health-of-families/index.html	52	52	37.55985117	52	0.5	52
https://oxford.shorthandstories.com/brain-language-literature/index.html	516	620	165.4933061	516	0.5988372	491
https://oxford.shorthandstories.com/brain-laurence-wainwright-recognising-the-impact-of-the-climate-on-mental-health/	52	52	181.9291181	52	0.3246753	26
https://oxford.shorthandstories.com/cancer-prevention/	181	465	246.6334625	336	0.4281768	284
https://oxford.shorthandstories.com/cancer-prevention/?utm_source=linkedin&utm_medium=referral&utm_campaign=oxfordshorthandstories	52	207	225.2459914	181	0.8563536	181
https://oxford.shorthandstories.com/cancer-prevention/?utm_source=twitter&utm_medium=referral&utm_campaign=oxfordshorthandstories	77	77	4.918536571	77	0.6623377	77
https://oxford.shorthandstories.com/dementia/index.html	2117	14692	1525.593469	3744	0.5752974	3718
https://oxford.shorthandstories.com/disease-and-degeneration/index.html	2014	2505	104.3341189	2221	0.6489493	2117
https://oxford.shorthandstories.com/epilepsy/index.html	826	955	185.6086184	775	0.6173121	775
https://oxford.shorthandstories.com/gender-and-sexual-identity/index.html	2453	2815	150.0123252	2556	0.4860659	2505
https://oxford.shorthandstories.com/innovation-dice/index.html	800	1291	131.9839491	1188	0.5738881	1188
https://oxford.shorthandstories.com/innovation-first-light-fusion/index.html	362	594	331.827765	362	0.6313646	362
https://oxford.shorthandstories.com/innovation-rollsroyce/index.html	52	77	6.88046589	77	0.6623377	77
https://oxford.shorthandstories.com/innovation-vaccine/index.html	258	258	77.98146937	258	0.1007752	258
https://oxford.shorthandstories.com/innovation-virtual-therapy-made-real/index.html	387	568	281.6327231	387	0.5784114	387
https://oxford.shorthandstories.com/maths-plus-cancer/	5216	8340	96.44022698	7540	0.6429345	7540
https://oxford.shorthandstories.com/maths-plus-cancer/?fbclid=PAAaYurshcEVNK4yeRk6R0EPFNZ6XY7ovPwiVS	52	52	0	52	1	52
https://oxford.shorthandstories.com/maths-plus-cancer/index.html?utm_source=linkedin&utm_medium=referral&utm_campaign=oxfordshorthandstories	129	155	18.72638979	155	0.7127072	155
https://oxford.shorthandstories.com/maths-plus-cancer/index.html?utm_source=twitter&utm_medium=referral&utm_campaign=oxfordshorthandstories	516	646	5.412371633	568	0.8703704	568
https://oxford.shorthandstories.com/mental-health-and-the-world-around-us/	3176	3899	109.316835	3357	0.6362405	3408
https://oxford.shorthandstories.com/mental-health-data/index.html	362	516	53.15414937	387	0.790224	413
https://oxford.shorthandstories.com/parental-and-maternal/	1239	1601	159.5551228	1239	0.6933167	1239
https://oxford.shorthandstories.com/professor-kj-patel/	3047	3899	147.8832023	3615	0.6644562	3615
https://oxford.shorthandstories.com/professor-kj-patel/index.html	129	232	244.1500896	155	0.6280193	207

Organise your content audit



Choosing how to **organise** your data

The recommended process for organising data depends on the size of your website:

Small websites

- fewer than 200 pages
- simple information architecture

Large websites

- more than 200 pages
- information architecture splits content into defined categories



Auditing large Oxford sites

When auditing large Oxford sites such as ox.ac.uk and sbs.ox.ac.uk, we recommend you split the results into a series of tabs that follow the information architecture (IA) of the sites.

stories.com/parental-and-maternal/	1239	1601	159.5551228	1239	0.6933167	1239
stories.com/professor-kj-patel/	3047	3899	147.8832023	3615	0.6644562	3615
stories.com/professor-kj-patel/index.html	129	232	244.1500896	155	0.6280193	207
stories.com/sleep-circadian-rhythm-and-mental-health/index.html	2324	2969	96.6112622	2505	0.6845778	2556
Home_category links	About	Admissions_undergraduate	Admissions_Graduate	Admissions_Continuing education	Research	News and events

It is recommended that content pages are recorded in tabs because:

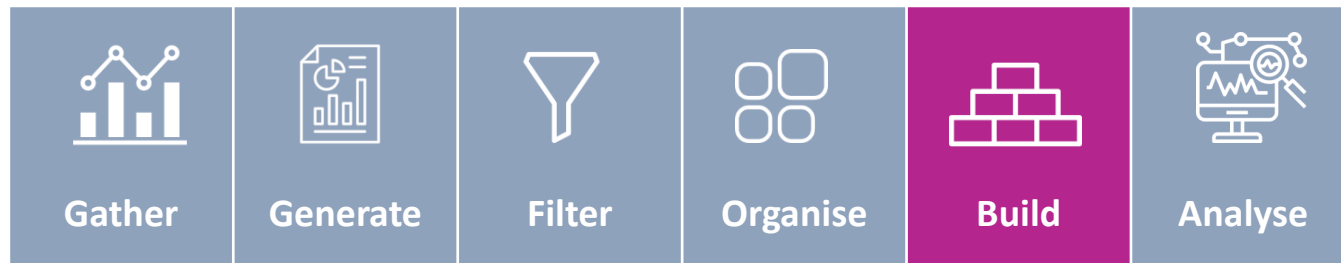
- the results for large sites are easier to view and analyse
- teams can more easily identify and review areas of content they are responsible for

Guidance for **auditing** sites of all sizes

It is recommended that the following content pages are recorded in separate tabs:

Top Level Content	Default URL content	News Content
<p>Any current and relevant content pages that sit on the top level of the site, but are not part of any category.</p>	<p>Content pages that sit on the top level of the site and are not assigned to any site category. 'Default URL' content pages appear to be outdated, redundant or published in error.</p>	<p>Any news content</p>
<p>These pages should be recorded separately, so that:</p> <ul style="list-style-type: none"> • content can be reviewed to see if it would better serve users if moved to an appropriate category • data can be analysed to judge if this content should be retained 	<p>It is recommended that 'Default URL' content is recorded separately, so that it is easier to identify and review potentially redundant pages before migration</p>	<p>All news content should be captured in a single tab. It will not be possible to migrate all news pages, and having a single list of news content is a good place to start when reviewing outdated pages for deletion</p>

Build your content audit



Using category tabs in your audit

On many Oxford sites, content is not tagged and most pages sit on the default site URL. In these cases, you choose to either:

Fill the category tabs manually

Cut and paste relevant data from the clean GA4 report into the audit, placing the data in the correct category tab. This can be a long process, especially if you are not familiar with the website, but the end results are easier to view and analyse

Record all content pages in a single tab

Cut and paste the clean data from the GA4 report into a single tab in the audit. This is a quick process, but it may be harder to interpret the data as there is no clear indication of where content sits on the website

Using category tabs in your audit

If you plan to record your website pages in category tabs, the process can be made easier by viewing page results in the GA4 report alphabetically.

For Oxford sites where content tagging is used effectively, you'll be able to quickly identify relevant rows and paste them into the correct audit tab.

For example, when building the Ox.ac.uk audit and populating the 'Students' category tab, viewing the GA4 report alphabetically grouped all content using the /students/ tag in the URL.

https://www.ox.ac.uk/students/graduation/transcripts
https://www.ox.ac.uk/students/graduation/transcripts/
https://www.ox.ac.uk/students/graduation/transcripts?wssl=1
https://www.ox.ac.uk/students/graduation/verification
https://www.ox.ac.uk/students/graduation/verification/
https://www.ox.ac.uk/students/graduation/verification?wssl=1
https://www.ox.ac.uk/students/graduation?wssl=1
https://www.ox.ac.uk/students/help
https://www.ox.ac.uk/students/help?wssl=1
https://www.ox.ac.uk/students/industrial-action/thanks
https://www.ox.ac.uk/students/life
https://www.ox.ac.uk/students/life/accommodation
https://www.ox.ac.uk/students/life/accommodation?wssl=1
https://www.ox.ac.uk/students/life/business-cards
https://www.ox.ac.uk/students/life/business-cards/order
https://www.ox.ac.uk/students/life/business-cards/order/confirmation
https://www.ox.ac.uk/students/life/clubs
https://www.ox.ac.uk/students/life/clubs/clubs
https://www.ox.ac.uk/students/life/clubs/clubs/about
https://www.ox.ac.uk/students/life/clubs/clubs/register-club
https://www.ox.ac.uk/students/life/clubs/clubs/registered-clubs
https://www.ox.ac.uk/students/life/clubs/clubs/registered-clubs/accounts
https://www.ox.ac.uk/students/life/clubs/clubs/registered-clubs/annual-re-registration
https://www.ox.ac.uk/students/life/clubs/clubs/registered-clubs/available-assistance
https://www.ox.ac.uk/students/life/clubs/clubs/registered-clubs/club-grants
https://www.ox.ac.uk/students/life/clubs/clubs/registered-clubs/committee-roles
https://www.ox.ac.uk/students/life/clubs/clubs/registered-clubs/current-clubs-and-societies
https://www.ox.ac.uk/students/life/clubs/clubs/registered-clubs/events-guidance
https://www.ox.ac.uk/students/life/clubs/clubs/registered-clubs/events-guidance/best-practice-guidance
https://www.ox.ac.uk/students/life/clubs/clubs/registered-clubs/events-guidance/health-and-safety
https://www.ox.ac.uk/students/life/clubs/clubs/registered-clubs/other
https://www.ox.ac.uk/students/life/clubs/clubs/registered-clubs/rules-regulations
https://www.ox.ac.uk/students/life/clubs/clubs/registered-clubs/rules-regulations/activities-overseas
https://www.ox.ac.uk/students/life/clubs/clubs/registered-clubs/rules-regulations/charity-status
https://www.ox.ac.uk/students/life/clubs/clubs/registered-clubs/rules-regulations/data-protection-act

Completing the Page and Link columns

Page column

For all 2023 Oxford content audits, the 'Page' column records the page address without the main site URL (for example, <http://www.ox.ac.uk/students/new/arrive> was recorded as </students/new/arrive>).

Full URLs can be batch converted into page titles by using the Excel 'find and replace' function

Link column

Always record the full URL in the 'Link' column and make the link active. You will be opening a lot of pages when working on the audit, so it's helpful to be able to open website pages in your browser quickly.

Completing the section column

Many categories on Oxford sites are split into sub-sections. For example, the ox.ac.uk 'Students' category is several sub-sections

Splitting categories into their component parts can be helpful when reviewing content.

But, if a category tab structure has been otherwise followed, splitting the 'Section' column into sub-categories is optional if resources are limited.

Sub-sections for the Student section

- Students/**new**
- Students/**academic matters**
- Students/**fees-funding**
- Students/**visa**
- Students/**life**
- Students/**welfare**
- Students/**graduation**
- Students/**news-list**

Common content types to record

Course information	Information	Listings	News	Form	Blog Page
Pages about specific courses available at Oxford (that is, content about individual courses, not about wider subject areas)	Any other page that publishes information for Oxford website users	Pages that link to a large volume of content – for example, news archive page or blog listing pages	News story pages	A digital form that users can complete and submit.	A blog page that sits within a website category (for example, research blogs published in the 'Research' category on Ox.ac.uk)

Note: Feel free to adapt the 'content types' to suit the content that's on your website

Completing the 'Who the content is for' columns

For each page, the 2023 Oxford audits include information on what audience the content is for.

When completing your audit, put a **Y** in each audience column the content is relevant to

- Staff
- Academics/researchers
- Students
- Prospective students
- General public
- Other specialised audience

Using the last modified column

Although data on when a page was most recently modified is one of the most helpful pieces of information when making decisions on content migration, this data is not always available.

If you are able to source a full list of site content from your in-house digital team, this may include 'last modified' date for each page, **but you will not get a 'last modified' date when running a GA4 report.**

Some Oxford sites include a publication date for all pages – particularly news content – and this should be recorded where possible.

But if no data is available, or you do not have the resources to inspect every website page for a published date, these cells should be left blank.

Completing the accessibility issues column

To gauge any accessibility issues that may pose challenges during migration, it's important to note any potential accessibility challenges.

Common accessibility issues include:

Illustrations

Make a note in the 'Accessibility issues' column of any illustrations published on your website that may pose a challenge for screen readers, or which could potentially be presented in accessible HTML (see appendix)

PDFs, Word and other Office Docs

Make a note in the 'Accessibility issues' column of any downloadable documents. Time permitting, open attachments to see if the content could be published in HTML and make a note of this in the 'Notes and recommendations' column (see appendix)

Note: You will not be able to capture all inaccessible documents as many are buried deep in content pages.

But auditing obvious downloads will give you a sense of any accessibility issues your site may need to consider

Analysing your audit



Completing the Notes & recommendations column

During migration planning, the 'notes and recommendations' column can be used to build a case for whether a content page should be migrated, reviewed or deleted.

The 'notes and recommendations' column is used to flag any key issues that should be considered during migration. The most common examples include:

- content that is part of a group of pages that get low views, and which potentially could be merged into a single, user-focused content page
- broken links on published pages or broken downloads
- content appears to be incorrect or outdated

Completing the migration challenge column

For each page, the migration challenge can be estimated and coded:

	Migration ready	Minor effort to migrate	Do not migrate
Content traffic	<ul style="list-style-type: none"> Significant page views (generally speaking around 300+ views in last 12 months) 	<ul style="list-style-type: none"> comparatively low page views (generally speaking under 300 page views in last 12 months) 	<ul style="list-style-type: none"> content was not viewed by users during the study period (zero views in last 12 months)
And/Or			
Content relevance	<ul style="list-style-type: none"> content appears to be in date and relevant 	<ul style="list-style-type: none"> content is possibly outdated page is one of multiple content pages about the same topic that could potentially be merged to reduce the page count 	<ul style="list-style-type: none"> content is clearly outdated (for example, an event or submission deadline has passed) content may be published in error

Thank you

Contact: elizabeth.mccarthy@admin.ox.ac.uk

November 2023

