Imperial College London
Projects

Environmental Research Group

UK Digital Air Quality Information Landscape Review

Independent analysis prepared by:

Andrew Grieve, Dr. Kayla Schulte, Adam Brighty & William Francis
Environmental Research Group, Imperial College London

November 2023

1 Executive summary

This review was commissioned by the Department for Environment, Food & Rural Affairs (Defra) to deliver insight into the current landscape of digitally available air quality (AQ) information in the UK. It is intended to assist the Air Quality Information Systems (AQIS) review panel, which is undertaking a comprehensive review of how government communicates AQ information.

The review examines the characteristics and evolution of air quality information channels (or AQ channels). Within the context of this review, AQ channels are "proliferating, internet-enabled AQ information sources that can blend personal information like geolocation, age and presence of pre-existing health conditions, alongside real-time (or near real-time) AQ estimates or measurements, to provide tailored recommendations for reducing exposure to pollution"[1,2].

Systematic searches of Google web search engine outputs were undertaken using a range of 22 AQ related search terms across 207 locations in the UK. Systematic searches of the Apple App Store and Google Play Store for mobile applications (apps) were undertaken using the same range of search terms. Systematic searches of social media platforms YouTube and X (formerly Twitter) were of limited value and are discussed further in the Discussion. Results were retained where they met the inclusion criteria; namely providing measured or modelled historic, real time or forecasted AQ data at a geographic location.

The results revealed 83 channels are available by way of websites, 83 channels are available via Apple apps, 34 channels are available as Android apps, and 3 channels from X (Figure 1). After removing duplicate channels (e.g., those that are available across two or more platforms), 146 unique channels remained. This formed the final list of channels used for this study. A clear increase in the number of channels offering AQ information was observed over the past 20 years (Figure 2).

The analysis of metadata gathered for these 146 channels revealed that multiple channels share common data providers, indicating a reliance on a common set of sources for constructing AQ information. Defra emerged as the most common data provider for AQ channels in the UK, followed by World Air Quality Index and Imperial College London's Environmental Research Group (Figure 6). Notably, some channels cited multiple data providers, particularly when offering forecasts alongside sensor or regulatory network data. It is important to note that data providers are not necessarily data producers. Of these 146 channels, 22 lacked clear attribution of the entities providing the underlying data for the information or messaging delivered by the channel.

Overall, the AQ information landscape in the UK has increased in size and complexity over the past 20 years with apps now rivalling websites as the main delivery platform (*Figure 2*) and an increase in the number of private companies entering the space. Key recommendations for consideration by Defra's AQIS review panel are summarised below.

Key recommendations:

- AQ information providers with the broadest reach have application programming interfaces (APIs). Defra should enhance its API offering and consider establishing an API to support the delivery of any new air quality messaging systems to maximise uptake.

- The digital AQ landscape consists of more channel providers and platforms for delivery than when the Daily Air Quality Index was updated in 2010. Defra will require a strategy of engagement with the new actors in the field to maximise adoption of any new system.

- Social media presents an enormous opportunity to connect with the public by providing an alerting platform as well as daily engagement. Defra should consider the development of social media strategy to harness this platform.

- Leveraging existing channels rather than creating new ones (i.e. creating new mobile apps) can optimize resources and streamline information dissemination.

- Defra should ensure consistent attribution and branding across all platforms to ensure data sources and messages are clearly recognised. Auto-delivered messages following API querying may be one option to achieve this.

Figure 1 Number of AQ channels accessible from key digital delivery platforms

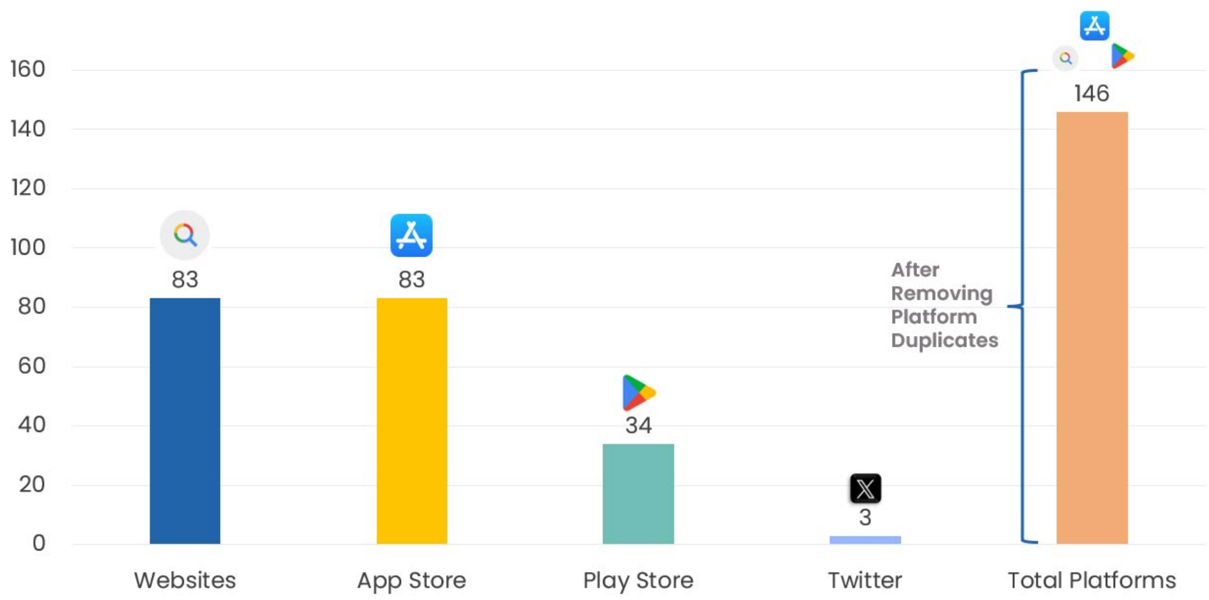


Figure 2 Unique AQ channels launch dates by platform

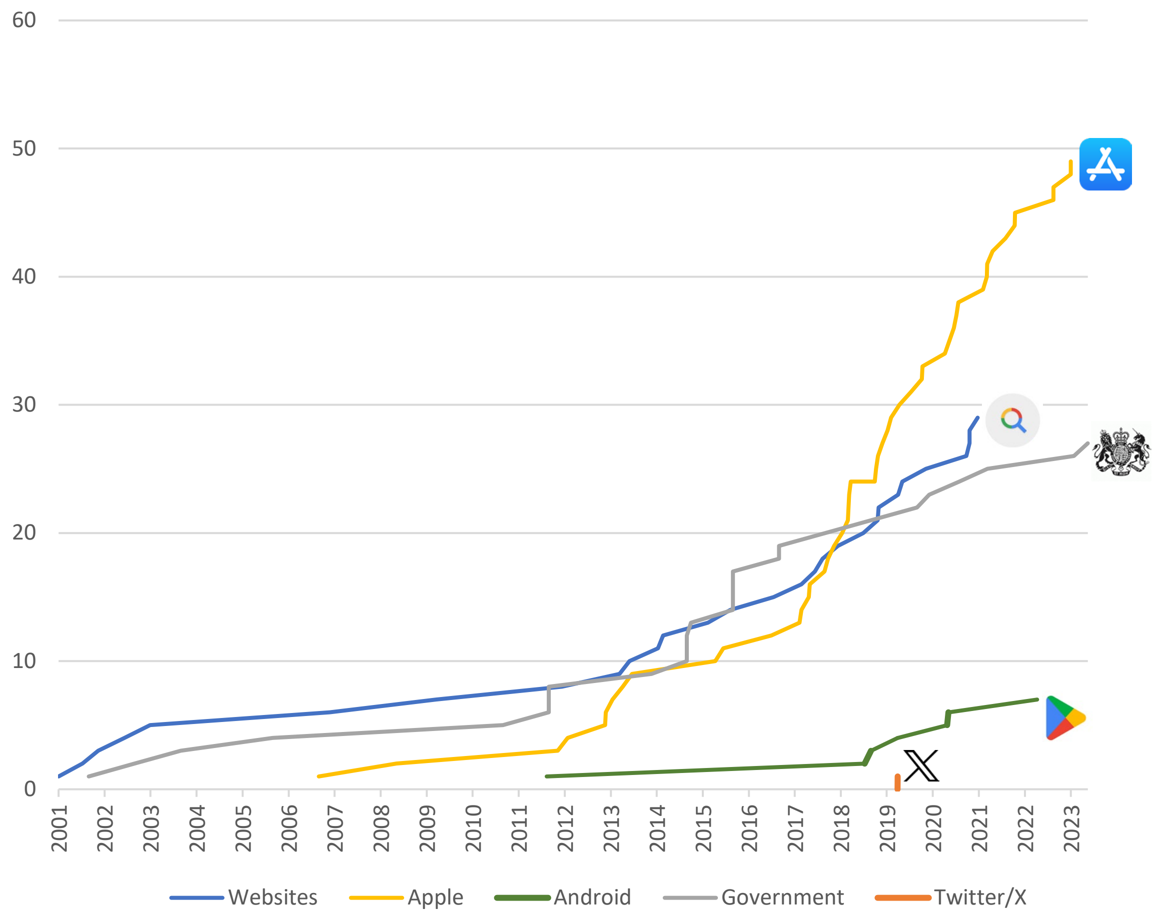


Table of Contents

List of Figures

List of Tables

2 Introduction

Availability and access to air quality (AQ) information has evolved significantly over the past two decades, from early reports of smog in newspapers to the proliferation of applications (apps), websites, social media platforms, and personal sensors that provide real-time AQ information. In parallel, the way people source AQ information has transformed, with a noticeable shift from the use of traditional media outlets such as newspapers and television to the Internet and social media.

Today, the digital AQ information landscape is broad and diverse, with multiple providers disseminating data and information across various digital channels. This diversity presents both opportunities and challenges when ensuring that the public receives accurate and up-to-date information. Within the context of this review, data is understood as measured or modelled air pollution concentrations communicated as pollutants or air quality index (AQI) summary values. Information is understood as the messaging and details delivered alongside air pollution data[1].

Despite a growing importance of AQ information[3], there is a notable absence of regular, systematic reporting on the current state of the AQ information landscape in the UK. This lack of structure hinders the government's ability to effectively communicate AQ issues to the public.

It is within this context that the Department for Environment, Food & Rural Affairs (Defra) is presently conducting a review of the Air Quality Information Services (AQIS) to enhance the government's communication strategies regarding AQ. To ensure that government efforts align effectively with the existing landscape, it is imperative to review and summarise the current state of AQ information provision in the UK.

This review seeks to bridge this critical gap by conducting a rapid review and analysis of the current AQ information landscape in the UK. By characterising the landscape, Defra can identify how best to interact with it in order to supporting the delivery of data and information that enables reductions in air pollution , while safeguarding human health and the environment.

To accomplish this, the review presents characteristics of the UK digital AQ information landscape, that is, AQ information sources that communicate about atmospheric conditions, i.e. measured or modelled air pollution emissions or concentrations relating to both:

- a specific geographic location, and
- a specific time (historic, near real-time, real-time, and forecast AQ conditions are included under this definition)

This review was conducted as a rapid response to Defra's call for evidence around the digital AQ information landscape, therefore a semi-automated approach towards data gathering and analysis was adopted. The details of this approach are described in the Methodology section.

The review describes the AQ information that is publicly accessible via digital channels and how the AQ information provided by Defra sits in the wider information communication landscape. It examines how the AQ data, which underpins existing AQ digital channels, is generated and by whom. Additionally, it reviews trends surrounding the prevalence of certain types of channels, changes in channels over time, and emerging trends across the information landscape. The review also examines channel specific differences in available information (e.g. geographic reach of data or type of information provided). Lastly, it reviews demographic and other socially rooted trends that may affect when, where, and how AQ information is accessed.

3 Research questions

The digital AQ information landscape is extensive and complex. To streamline our research efforts and deliver a rapid review of this landscape, we identified a set of primary and secondary research questions to guide this inquiry.

3.1 Primary research questions

1. What can be said about the features and characteristics of the current digital AQ information landscape and its underlying systems?
2. What can be said about the data origins of the current AQ digital information channels?
3. What are the trends in reach / access to AQ digital information channels – including any insights that can be observed through available metadata/analytics on where, how, and when different groups access AQ information?
4. Are there any channel specific differences in digital data provision (for example by geographic location or information type)

3.2 Secondary research questions

1. How has the UK digital AQ information landscape changed in recent years?
2. What technological developments may have impacted changes and developments in digital communication over recent years.
3. Are there any emerging trends in digital information channels relevant to future information dissemination across this landscape?

4 Methodology

4.1 Approach

This project employs a mixed methods approach (systematic search and metadata analysis) to gather the appropriate data to effectively describe and map the UK AQ information landscape. This approach builds on methods developed in *Schulte, 2022* for conducting a systematic technology review to identify and gather data about publicly available AQ channels[4].

Alongside the widespread use of internet-connected technologies such as AQ channels, there has been an increase 'digital trace data' such as app download data or Domain Name System (DNS) server records[3]. These data can help contextualise and characterise aspects of the digital AQ information landscape, while taking care in their interpretation to not conflate with strictly positivist assumptions associated with the notion of 'raw data'[5]. That is, "data are not simply natural and essential elements that are taken from the world in neutral and objective ways"[26]. The methodological approach adopted in this review considers both air pollution data and digital trace, user analytics data as entities that are produced through complex sociotechnical systems[6,7]. Acknowledging this is important to ensure results from this review are not extrapolated out of context.

4.1.1 Search protocol

Searches of common and relevant digital delivery platforms[8,9], including Google, Apple App Store (iOS) and Google Play Store (Android), were conducted using search engine query software SerpAPI[10]. Leveraging SerpAPI allowed for an efficient, comprehensive, and replicable search of each digital delivery platform. SerpAPI was used to emulate a selection of devices including Apple iPhone mobile device, Android mobile device, as well as standard Windows and Mac desktop devices. Combined, iOS and Android account for 99% of the smartphone market in the UK[11]. Google search represents 93.7% of the UK market share[12].

Barriers to access restricted the evaluation of AQ information contained on specific social media platforms[13]. Facebook & Instagram (Meta) provide a research portal (FORT) however, the lengthy signup and vetting process meant we were not able to include this in this rapid review. Tik-Tok's research portal is only open to academics based in the US.

SerpAPI was used to conduct a search of YouTube. Unlike Google search, it was not possible to set the geographic location of searches in YouTube. This seemed of limited importance however as searches performed using SerpAPI and manually with phones/laptops for the search terms returned the same results.

Twitter/X API is provided at 4 levels: free / basic / pro and enterprise. Only the Pro level allows querying of usernames/IDs. The ability to search Twitter for users based on a search term is not available in Twitter's current API. Instead, researchers used Twitter's advanced search function and manually extracted the top 20 accounts returned for each search parameter. These were then rationalised to unique accounts and searched manually to determine if they were AQ channels.

Radio and TV are important broadcast channels in the UK. However, there is no easily searchable archive of radio and TV programmes that allowed for mining of information to determine where and when air pollution has been included in news or weather bulletins. For this reason, broadcast media was not included in this search.

There are also examples of outdoor digital displays or billboards, such as the emergency alerts issued at bus and tube stops during high pollution events by the Greater London Authority (GLA)[14], and campaigns showcasing live data including those by AddressPollution[15] and Global Action Plan's 'Breathable Billboard'[16]. Identifying these cases across the UK requires a different methodological approach and was therefore beyond the scope this review. The use of outdoor displays, radio and TV are revsited in the Discussion and Conclusion of the review.

4.1.2 Search parameters

A list of search terms was obtained from *Schulte, 2022* to help generate an exhaustive list of publicly available AQ channels serving individuals in the UK. We conducted a test to determine whether geographic location (IP address) had an impact on search results. The researchers manually set the browser location to two public library locations, one in Newcastle and one in Bristol. Outcomes of the test revealed that city-specific search results were returned.

To address the research questions most directly and Defra's areas of interest, searches were emulated at both the city, county, and shire level. This is the ideal geographic scale to capture area-specific content returned to an individual searching for real-time AQ information at any given location across the UK. It reflects the geographic resolution at which Google and App store searches are returned and also ensures that the resulting database does not solely include results for urban areas.

The list of UK cities was taken from the UK Government official webpages[17] whereas the list of UK counties was formed upon consultation of a wider range of resources. The number of UK counties has changed significantly over time and so historical counties were removed from the locations list:

- In England, only current ceremonial counties were included.
- In Scotland, registration counties were used instead of Lieutenancies. This makes little difference, apart from the cities of Aberdeen and Edinburgh not being included as counties. The City of Glasgow was also excluded from the county list as SerpAPI was not able to read it as a location. These cities were included in the city list.
- In Northern Ireland, all counties were included except for the cities of Belfast and Derry for the same reasons as the Scottish cities.
- In Wales, 22 counties and county boroughs have been used to divide Wales since 1996, due to the Local Government Act of 1994. The county boroughs do not fall within other counties and so have been included within the counties list along with the remaining Welsh counties.

In 49 cases, SerpAPI could not retrieve data from the location as they are not included within SerpAPI's 'supported location list'. In these cases, postcodes that matched the county town or the London borough were used instead of location name. Only one postcode was chosen for each London borough to prevent repeated results. Only four locations, all minor Scottish counties, were not found within the supported location list. A total of 207 locations were included, with searches repeated for each 'search term'. For a complete list of search terms and locations, see Appendix 8.1 and 8.2 respectively.

4.2 Data

When conducting searches via the Google search engine, a maximum of 100 results per search term were retained. For context, each Google search page returns 8 – 10 results[18]. The high volume of retained results ensures the capture of all possible AQ channels returned to a given user.

The same list of search terms was used for obtaining results from the Apple App Store and Google Play Store. It is not possible to change search location across the two app stores, and any search term that included the word 'nowcast' returned no results. SerpApi was not able to consistently return the same relevant information for the Apple and Google Play app stores (Appendix 8.3.1).

Reproducible scripts containing all search parameters can be accessed via Github. The process for refining the search results is described in the following section.

4.2.1 Inclusion and exclusion criteria for AQ channels

The criteria delineating which AQ channels are retained for analysis establish the necessary boundaries for the data collection exercises undertaken in this review.

Table 1 AQ channel inclusion and exclusion criteria for landscape review

Inclusion Criteria	Exclusion Criteria
• Publicly available[i] • Digitally delivered (this includes websites, mobile apps, social media channels) • Delivers historic, real-time, or forecasted (or any combination thereof) AQ information for a given geographic location (maximum regional level) at a minimum annual[ii] temporal frequency • Includes information available in England, Scotland, Wales, or Northern Ireland	• Non-digital channels (such as leaflets) • Channels that only provide pollen counts • Channels that only provide weather data • Reports such as government reports that summarise historic AQ data.

This review adopted a 'three tier' refinement process to systematically retain relevant search results. The first tier involves filtering the SerpAPI results to produce a dataset with unique entries, thus avoiding any AQ channel duplicates.

The second tier leverages a 'crosswalk', or dataset of 50 common terms contained in descriptions of AQ channels meeting the criteria above. Descriptions returned for each of the apps, along with html text for each Google search results were examined to see how many terms matched the crosswalk. A score was generated for each unique search result and used to aid in the process of determining which results are most likely to meet the criteria for inclusion as an AQ channel. Details of this process are described in Appendix 9.6.1.

The third tier incorporated manual checks to review each entry or row in the dataset and to confirm that the AQ channel in question meets the following criteria:

1. Is it digitally available?
2. Does it provide air pollution data (not just pollen or weather information)?

[i] Requires access to an internet-connected device and is available to the UK general population.
[ii] Inclusion of channels that supply annual averages allows the report to capture how the digital AQ landscape includes "repackaged" air pollution data contained in government reports.

3. Does it provide additional data beyond simply summarising historic data (i.e. government reports, PDFs, etc)?
4. Does it provide data for UK?

If the responses to all four questions are ‘Yes’, then the entry is retained. The remaining entries are then populated with metadata.

4.2.2 Metadata categories

To evaluate the reach of AQ channels emerging from the systematic technology review, metadata was gathered manually for each channel retained throughout the review. A complete list of the metadata fields can be found in Appendix 9.3.

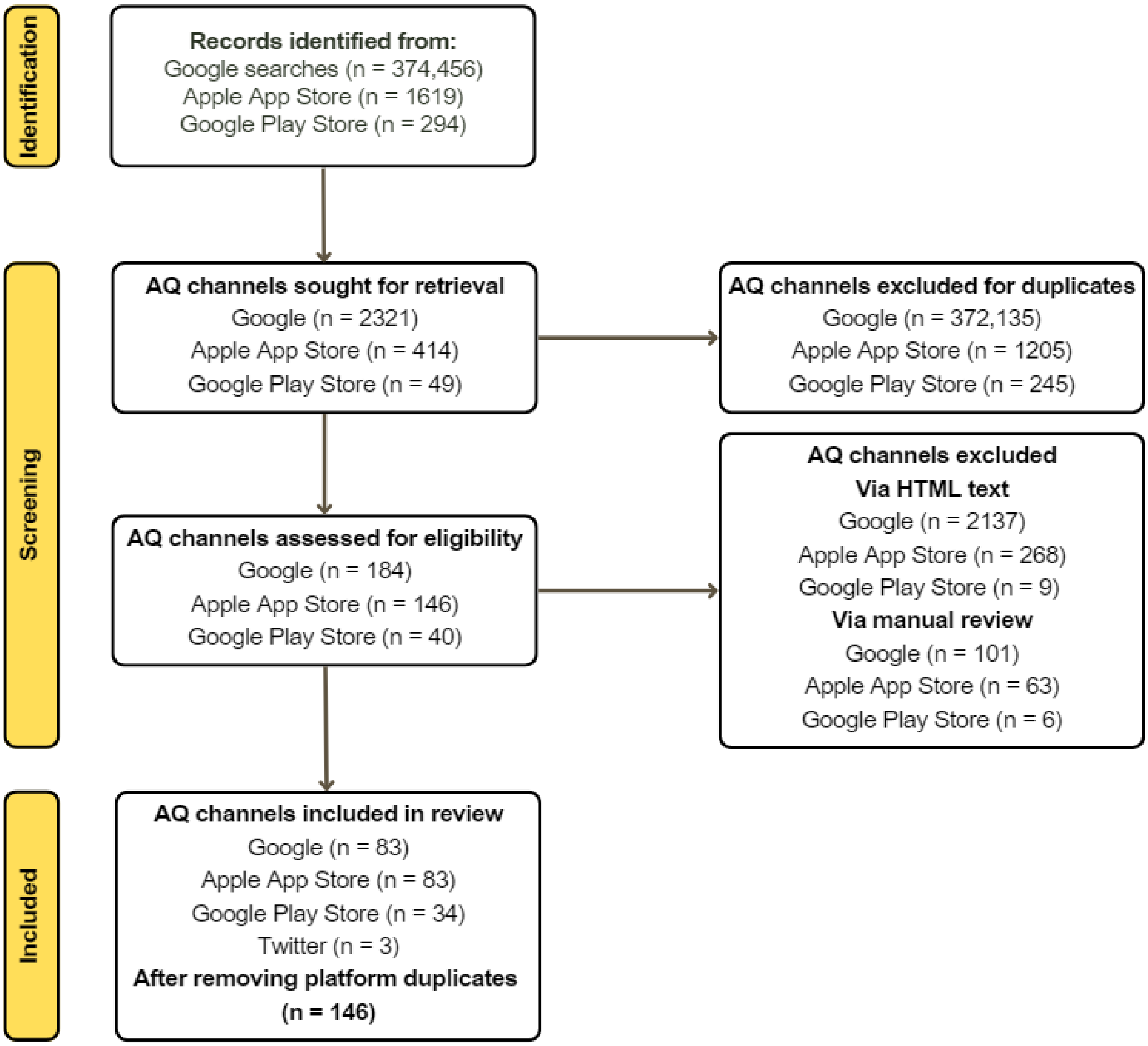
We adopted a partially computational approach for gathering metadata. Several of the high-level metadata fields listed in Appendix 9.3 were returned as a part of SerpAPI querying process. The remainder of the fields were obtained by matching key terms with the text from the app’s description, or html text obtained from the homepages of URLs returned via the Google search. The remaining metadata fields that remained unpopulated following the computational procedure were manually searched and input into the dataset.

Channel launch date was determined using a range of methods. For websites, in most cases website registration date was used (using whois.com). Alternatively, the data created was extracted from the page source. If neither of these were available, the earliest data available on the site was used. For apps, SerpAPI returned app launch dates.

4.2.3 Framework for analysis

A structured content analysis of the resulting database of AQ channels was employed to respond to both the primary and secondary research questions. Platforms for data collection and analysis included RStudio, SerpAPI and the Twitter API. R Studio packages used include: tidyverse, jsonlite, readxl, xlsx, httr2, httr, lubridate, rvest, data.table, urltools, rlist, r.utils, RSelenium, and XML.

Figure 3 Refinement of systematic technology search results



5 Results

The systematic search returned 376,369 items, which was refined to 146 AQ channels that met the criteria for inclusion. Of the 146, 83 channels were available via websites, 83 were available as Apple apps and 34 as android apps, with certain channels available across all three platforms (Figure 3). Results of the metadata excavation exercise revealed features and characteristics of the current digital AQ information landscape, their underlying systems, origins of the data presented, while highlighting emerging trends and how the landscape has evolved over time.

Metadata pertaining to number of downloads indicative of reach was available for a sample of the 146 channels. It is necessary to note that download counts represent an approximate figure and correspond to total downloads globally.

5.1 Evolution of the digital UK AQ landscape

The number of AQ channels has increased over time and the current dominant delivery platform are mobile apps (Figures 4 - 7). The number of App launches overtook the number of website based AQ channel launches in 2018. The results show a clear increase in the number of AQ channels from 2001 to 2023, with upticks in 2001, 2012, 2015, 2019 and 2021 (Figure 4).

Figure 4 AQ channels launches over time

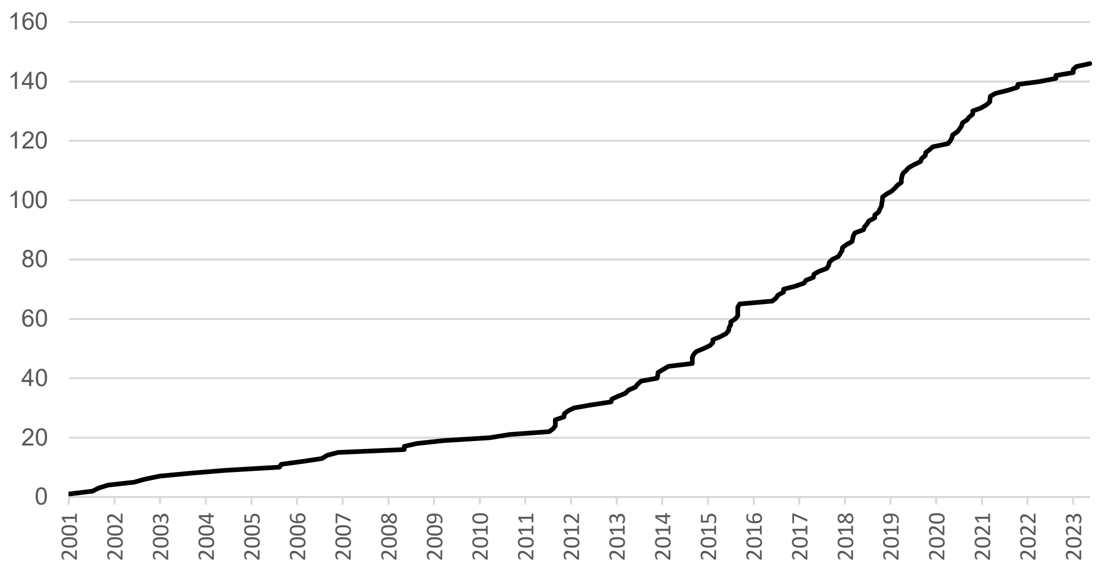
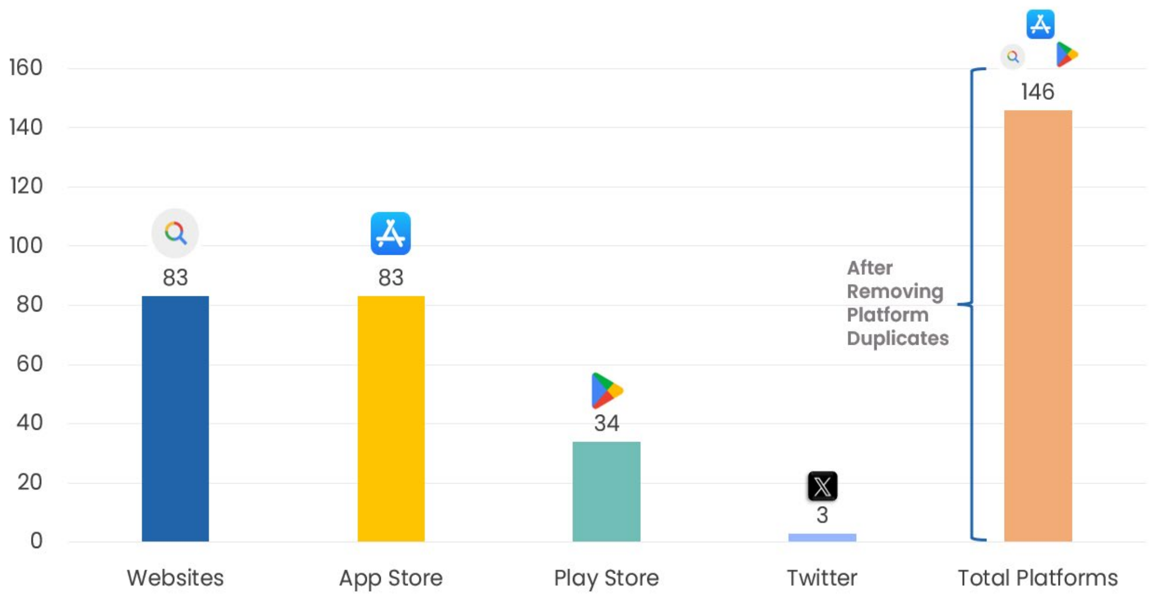


Figure 5 Number of AQ channels accessible from key digital platforms



5.2 AQ channel data providers and data producers

Analysis of the metadata for each channel demonstrated that multiple channels were associated with the same 'data providers.' This means that the data used to construct the information or messaging delivered across multiple channels is derived from a common set of data providers (Figures 9, 10, 12 - 17). Certain channels cite multiple data providers. This is often the case if channels provide a forecast in addition to sensor or regulatory monitor data. For example, the channel 'AirMatters' obtains data from one provider to present a real-time air pollution measurement for a given location, while obtaining data from an additional provider to present a forecasted (modelled) air pollution concentration. The most common data providers across the 146 AQ channels was Defra, followed by WAQI, and then Imperial ERG.

It is important to note that data providers are not necessarily 'data producers'. For example, World Air Quality Index (WAQI) collates data from different data producers around the world and makes them accessible through their own API. On the other hand, Imperial ERG is both a data provider and a producer. That is, ERG produces multiple air pollution datasets (i.e. Breathe London, LondonAir nowcast, LondonAir monitoring network) and makes them available via their own API, websites and apps. A channel may list Imperial ERG as a provider but may only be obtaining one of ERG's AQ datasets. Local councils may collate data from different data producers (Figure 17).

Figure 6 Growth of individual AQ data providers over time (largest 5 named)

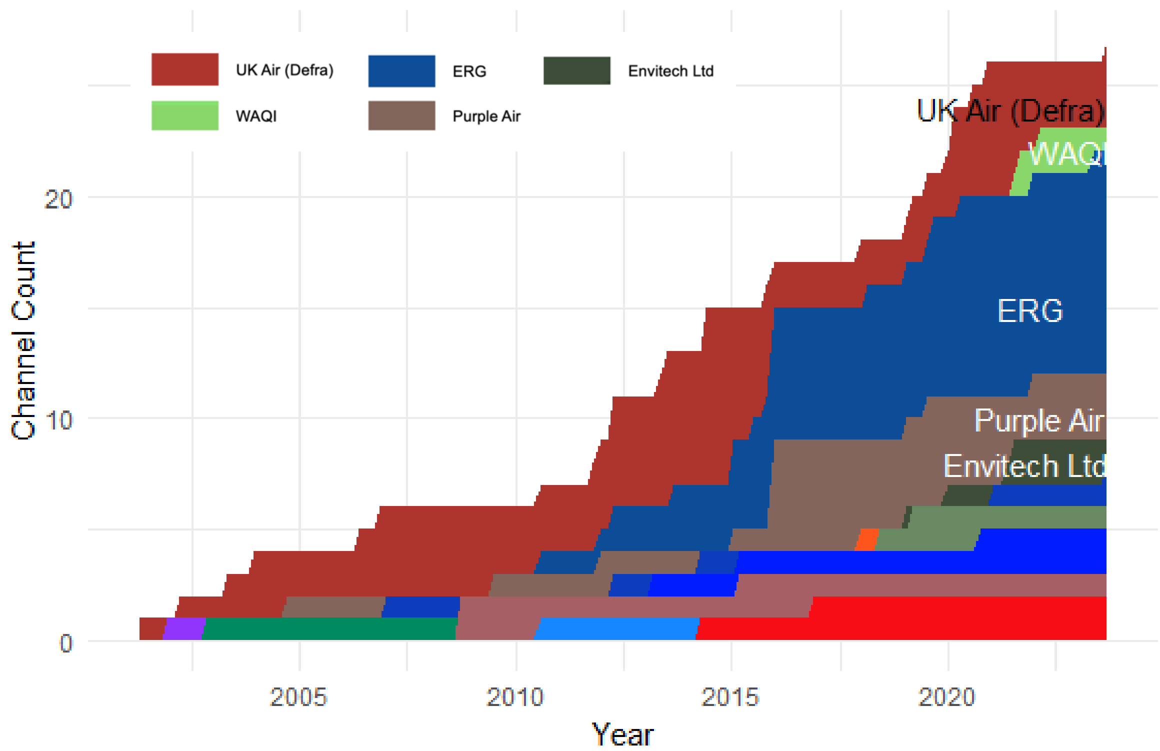


Figure 7 Growth of total number AQ data providers over time

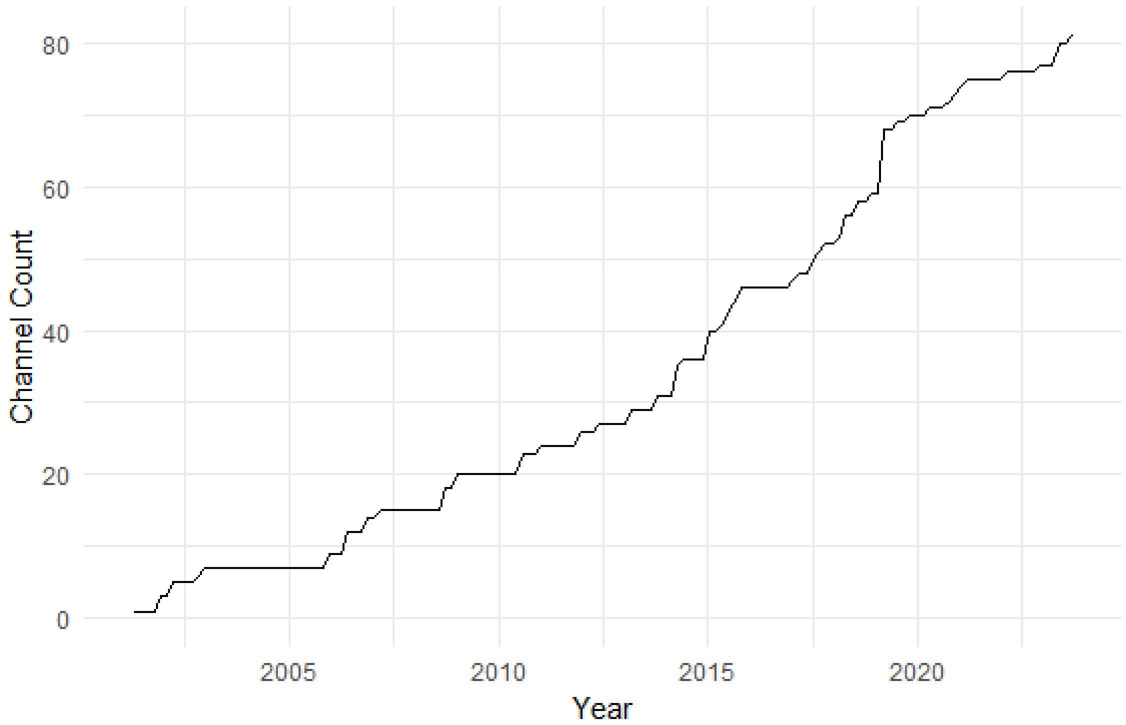


Figure 8 Largest AQ data providers

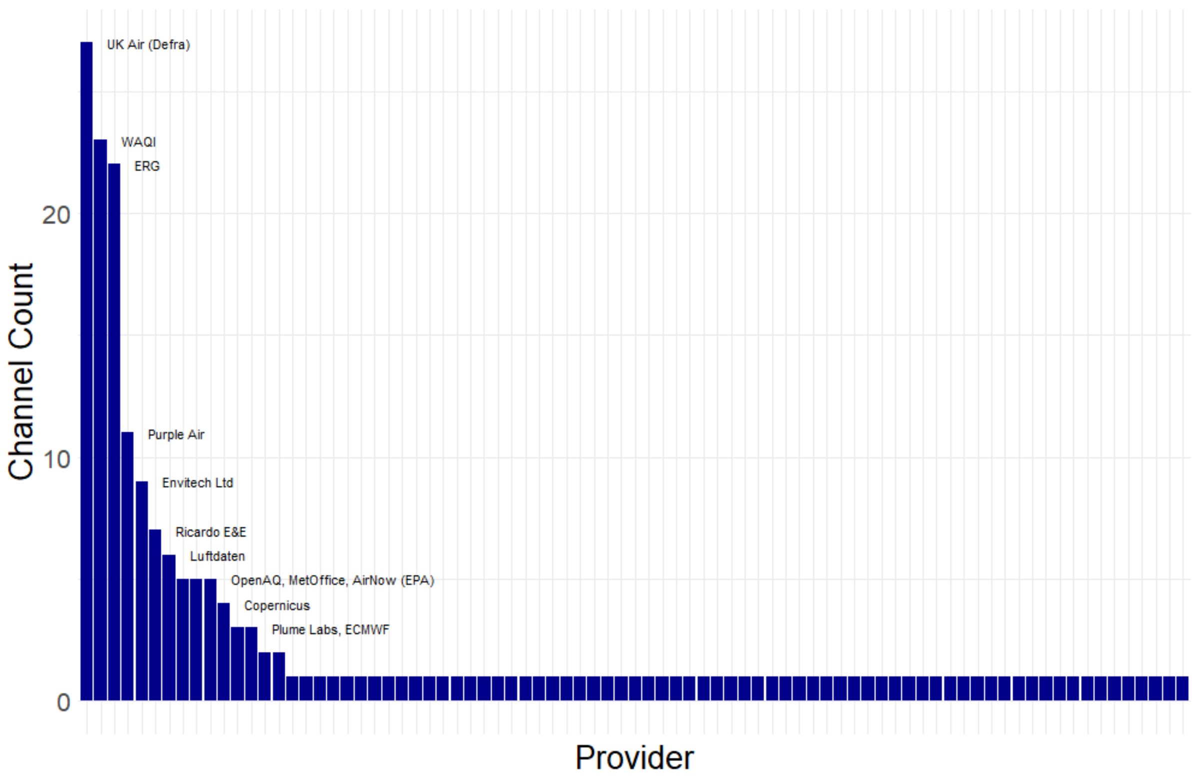
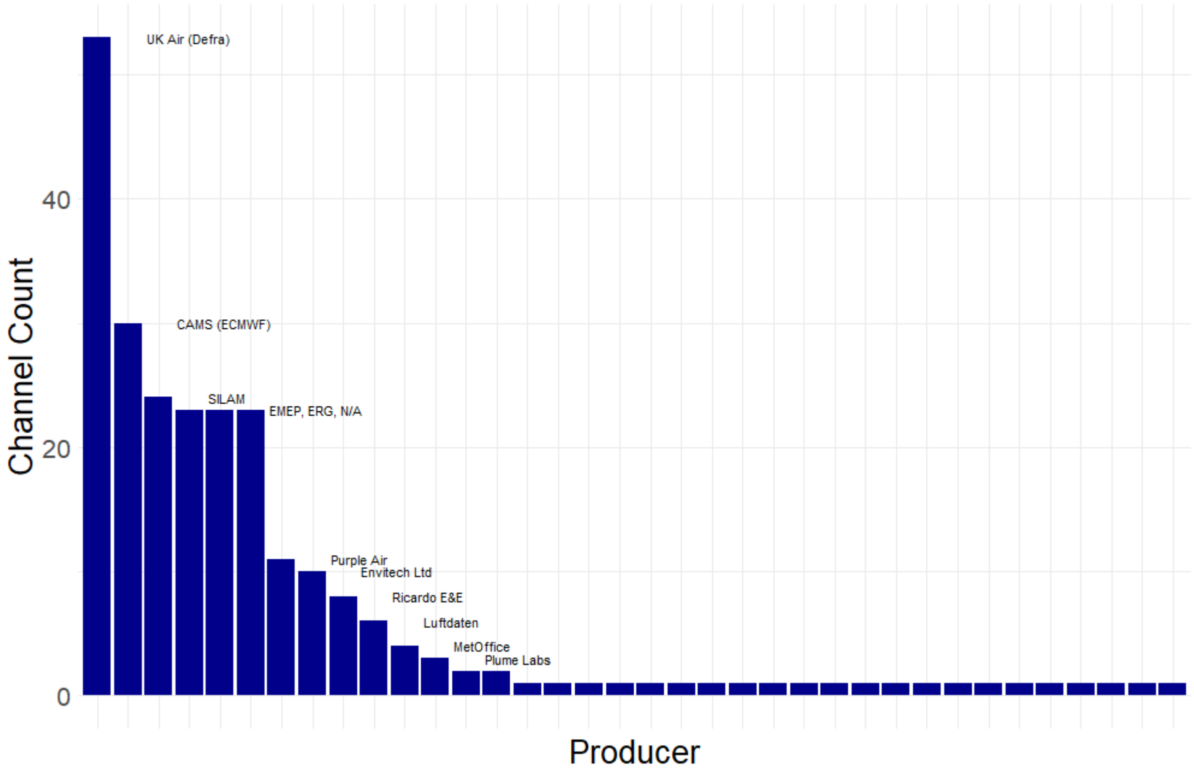


Figure 9 Largest AQ data producers



5.2.1 AQ data pathways

The following section contains dendrograms illustrating how the data from core data producers flows into data providers and ultimately into channels presented to the public. Figure 11 below describes the structure of the dendrograms.

Figure 10 Dendrogram structure

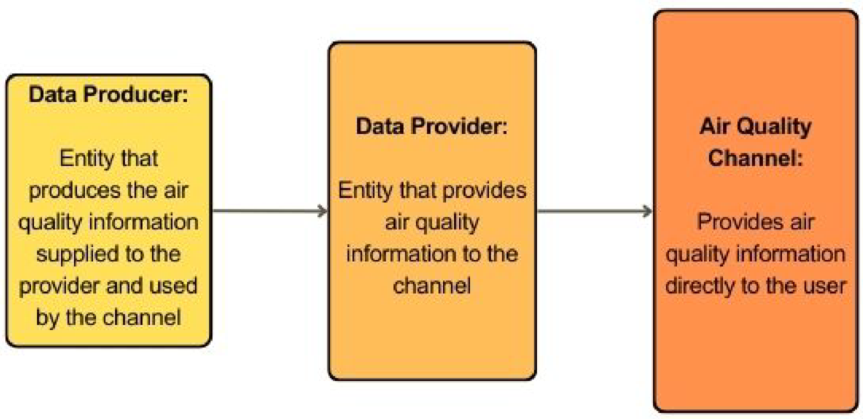


Figure 11 Dendrogram showing UK-Air (Defra) data pathways

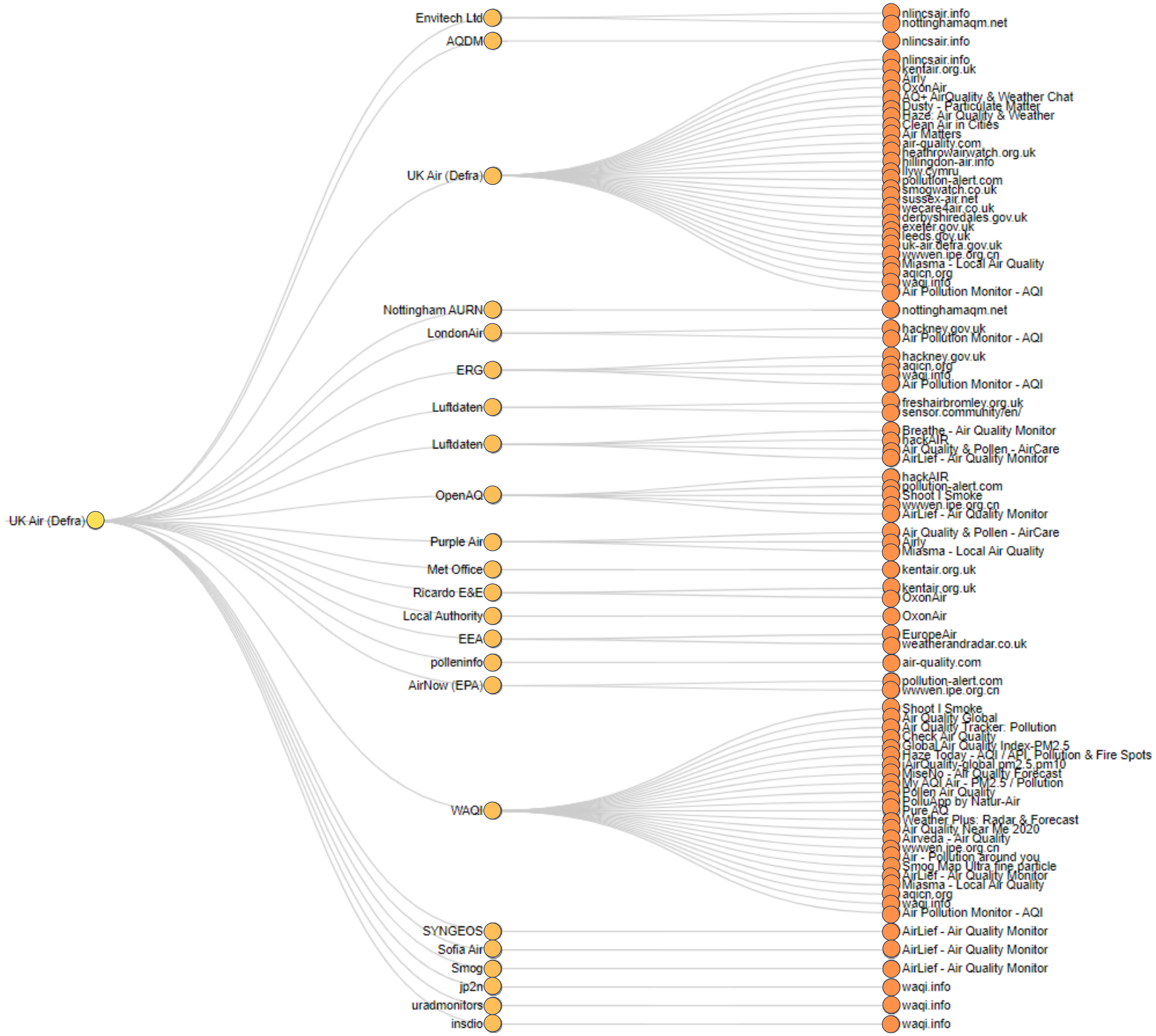


Figure 12 Dendrogram showing Copernicus Atmosphere Monitoring Service (CAMS) data pathways

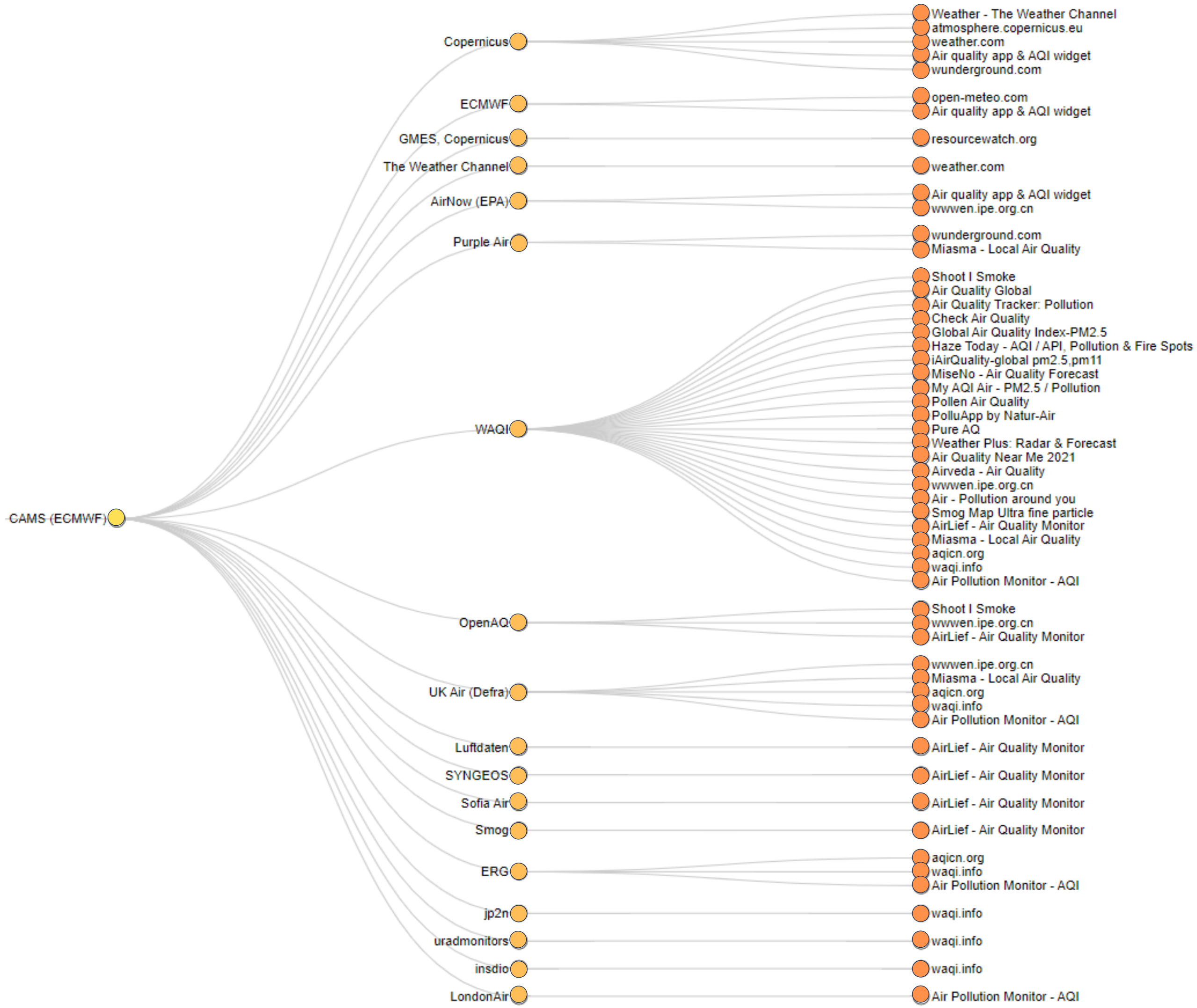


Figure 13 Dendrogram showing Environmental Research Group (ERG) and System for Integrated Modelling of Atmospheric Composition (SILAM) data pathways

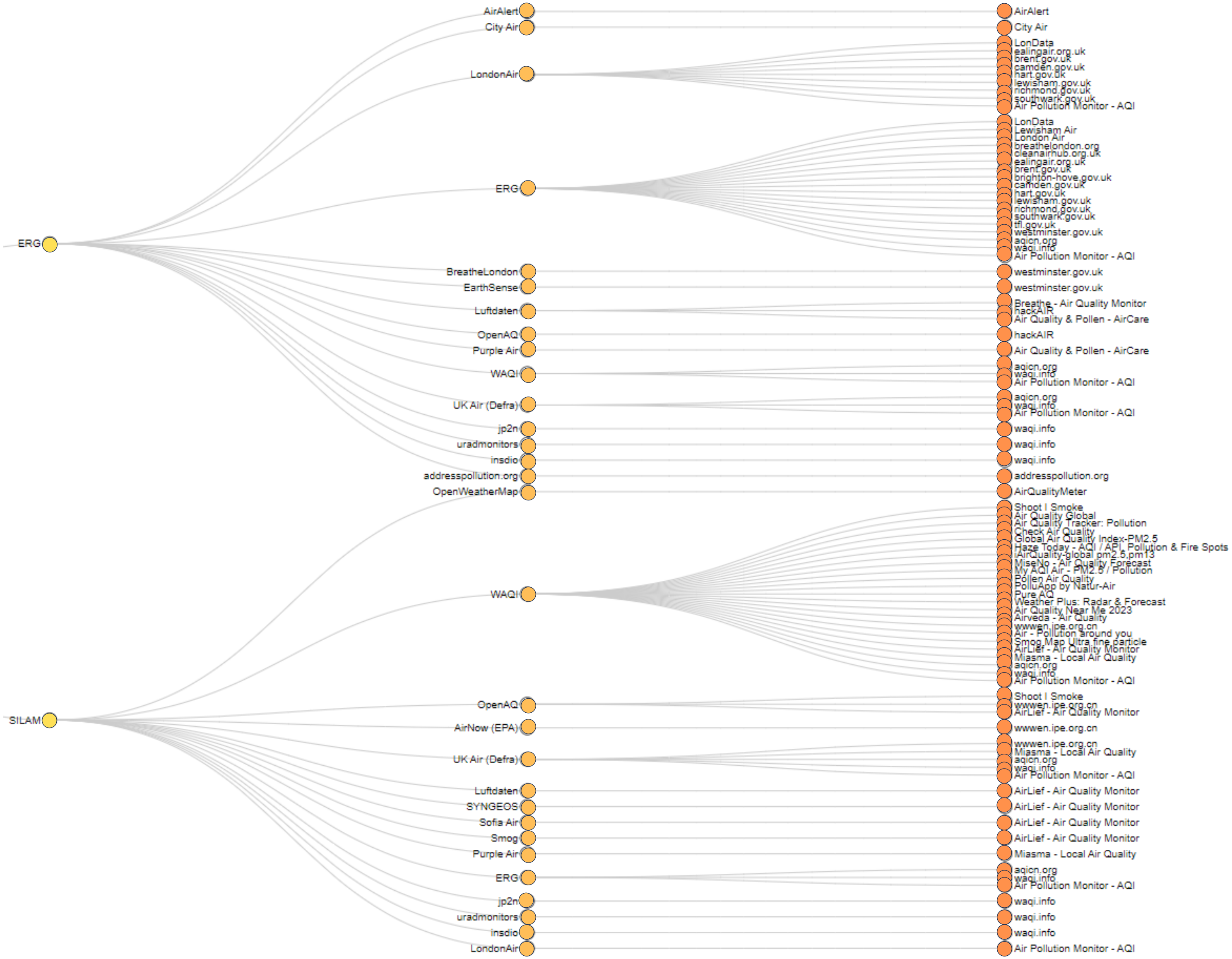


Figure 14 Dendrogram showing the European Monitoring and Evaluation Programme (EMEP) and PurpleAir data pathways

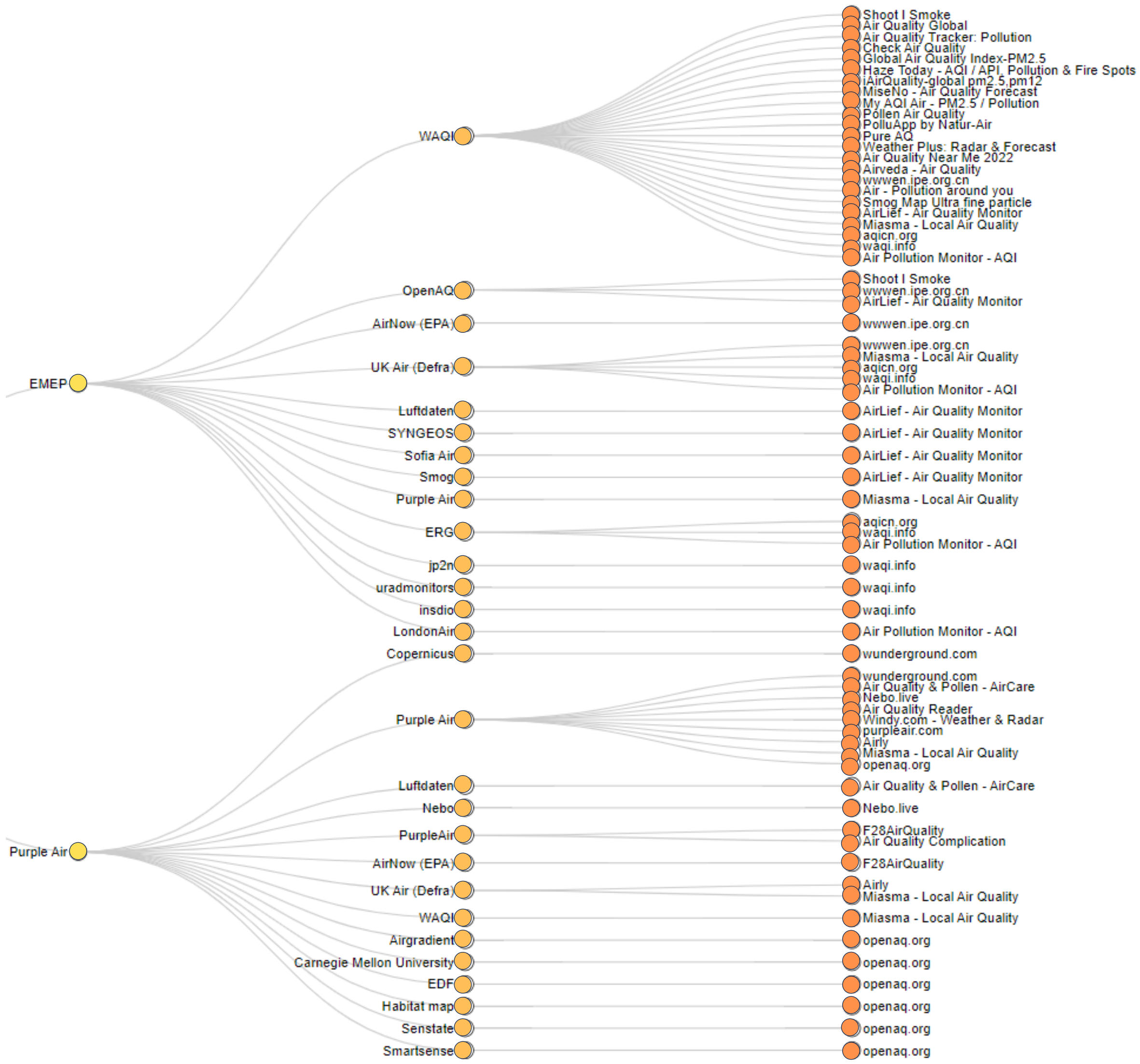


Figure 15 Dendrogram showing Envitech, Lufdaten, Ricardo Energy & Environment data pathways and those where no data provider was noted

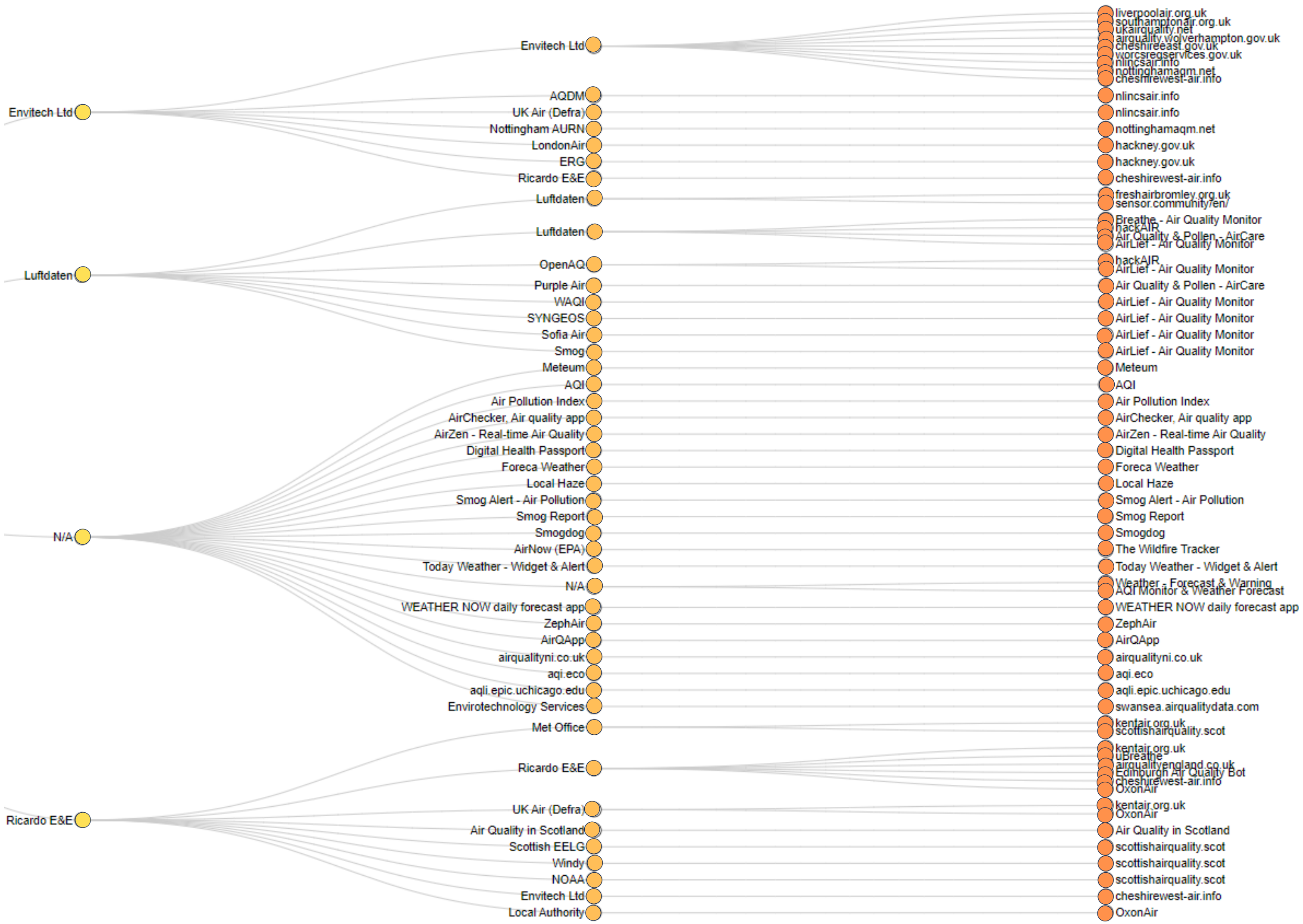
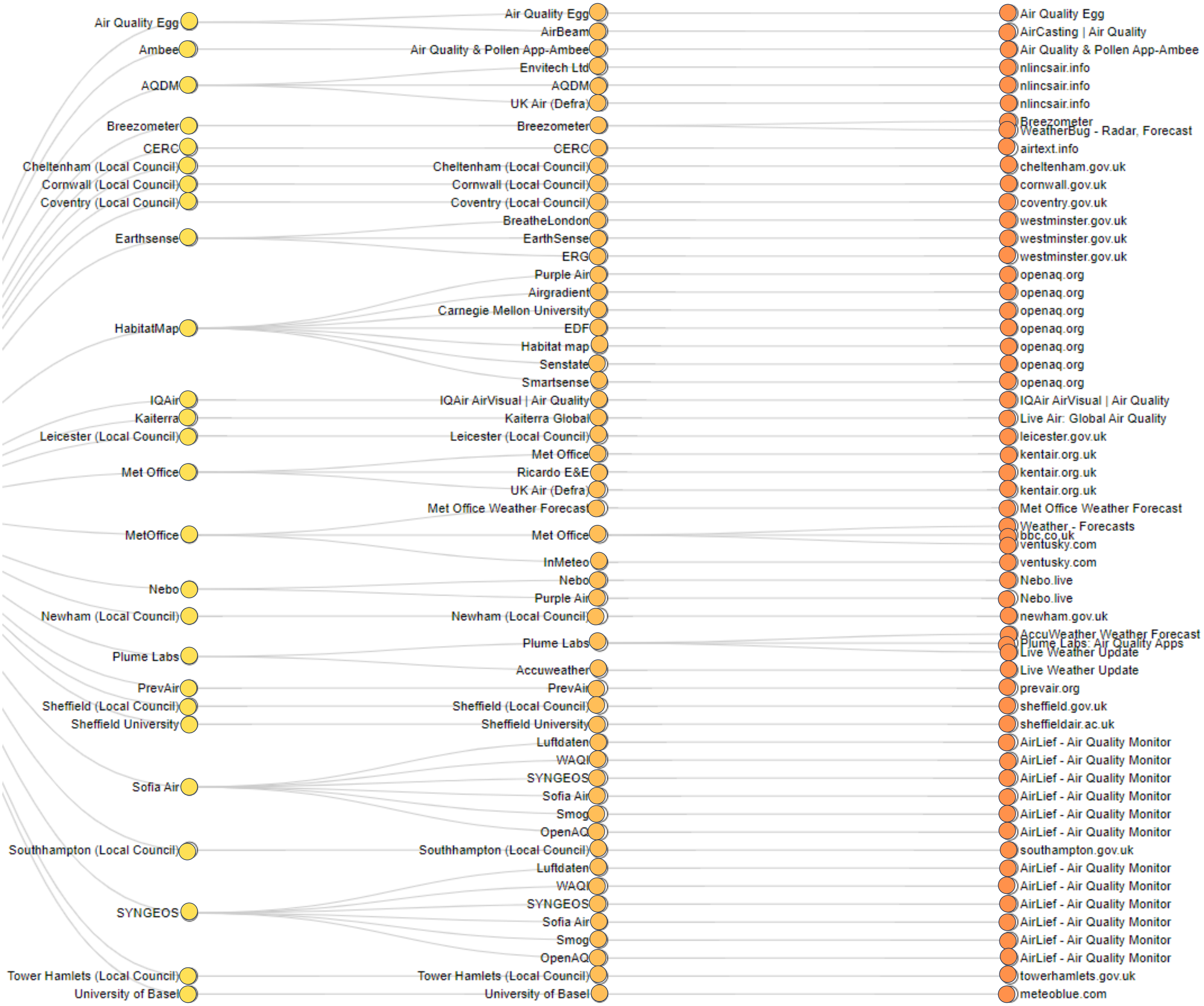
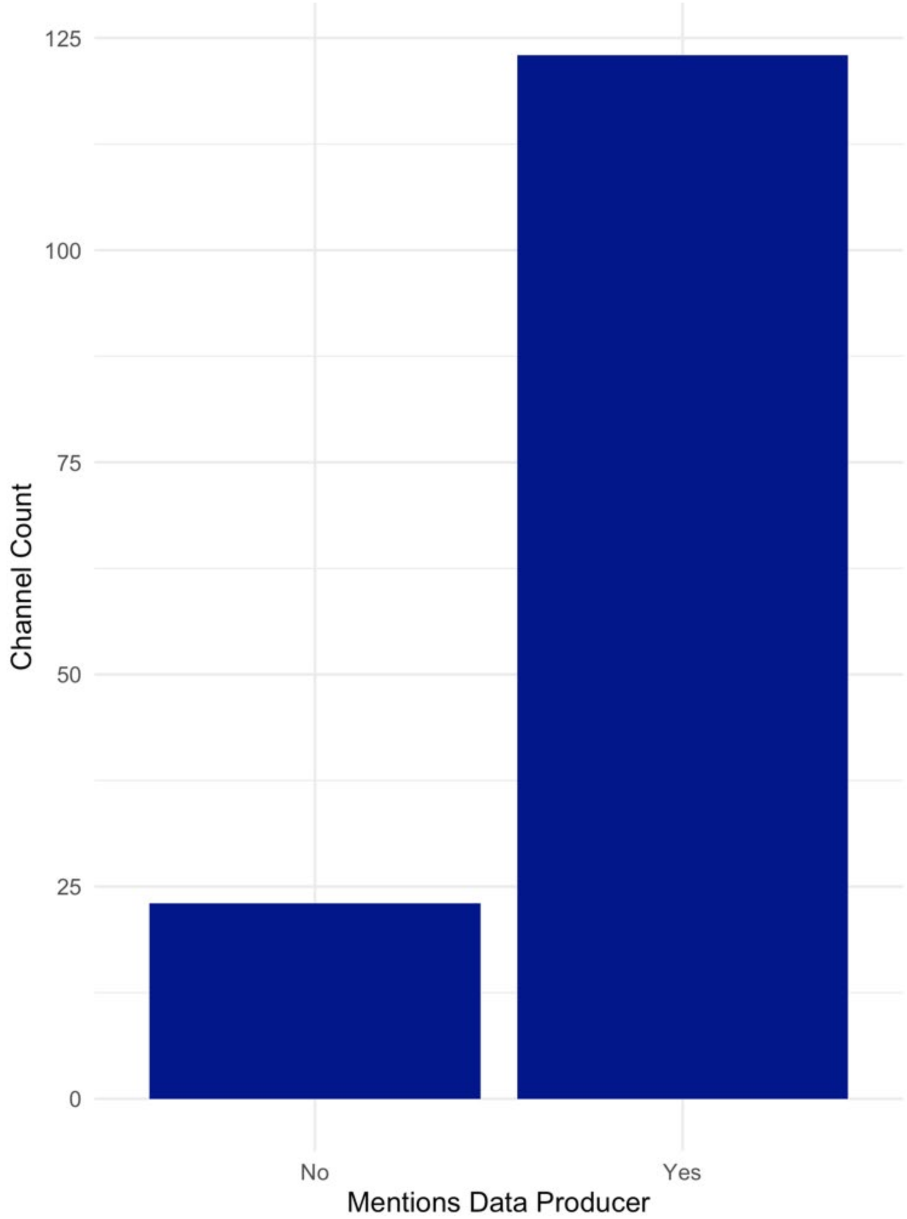


Figure 16 Dendrogram showing data producers feeding 8 of fewer channels



There were also 22 instances where AQ channels did not provide a clear indication as to who was providing the data underpinning the information or messaging delivered. In these cases, even if the characteristics of the channel data had features similar to those of specific data providers, no data provider was listed (Figure 16 and Figure 18).

Figure 17 AQ channels count where no data provider was cited



Analysis of the metadata demonstrated that AQ channels across the UK digital AQ information landscape derive data predominantly from regulatory monitors (Figure 19). Over time, the split of data sources fortifying AQ channels has expanded to include a greater proportion of sensor-based data sources (Figure 19). However, monitoring data maintains the largest data source underpinning AQ channels (Figure 19).

The results show that the majority of channels do not provide an AQ forecast (Figure 20). Of the channels which do provide a forecast, 1 day ahead was the most common forecasted, followed by 5 days (Figure 21).

Figure 18 Digital Air Quality Channel data sources over time

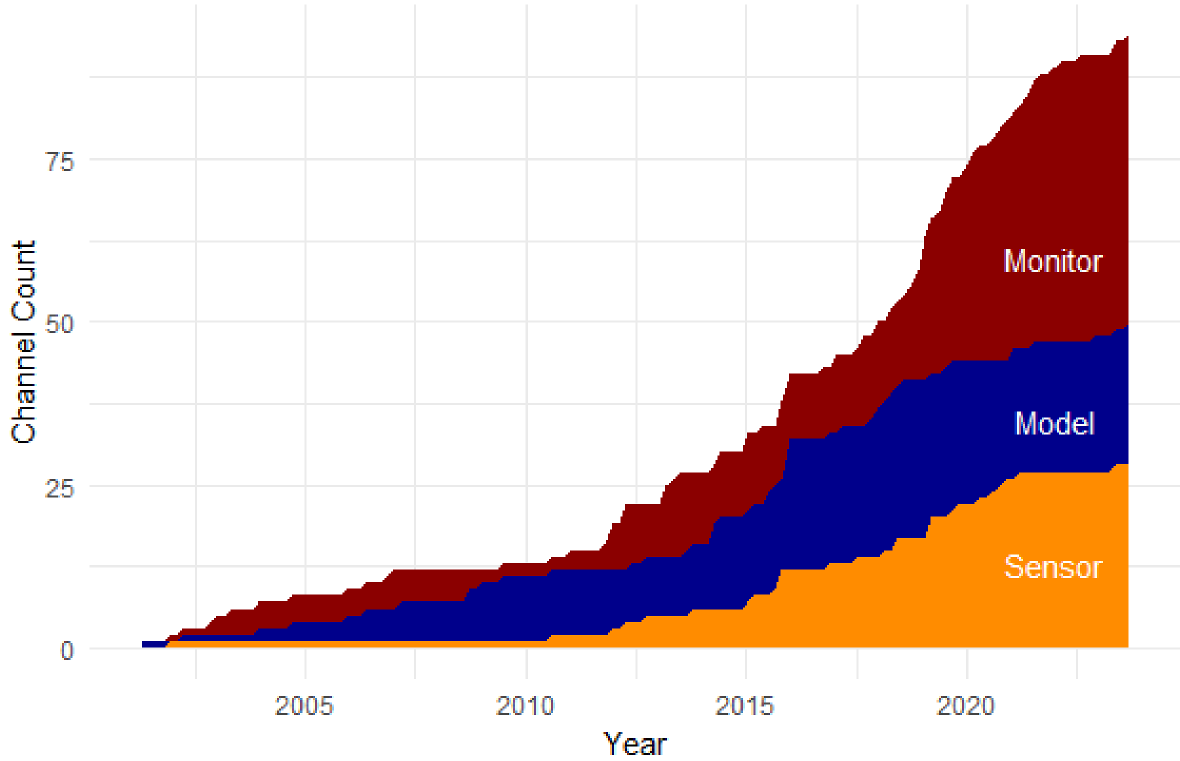


Figure 19 Number of AQ Channels providing an air pollution forecast

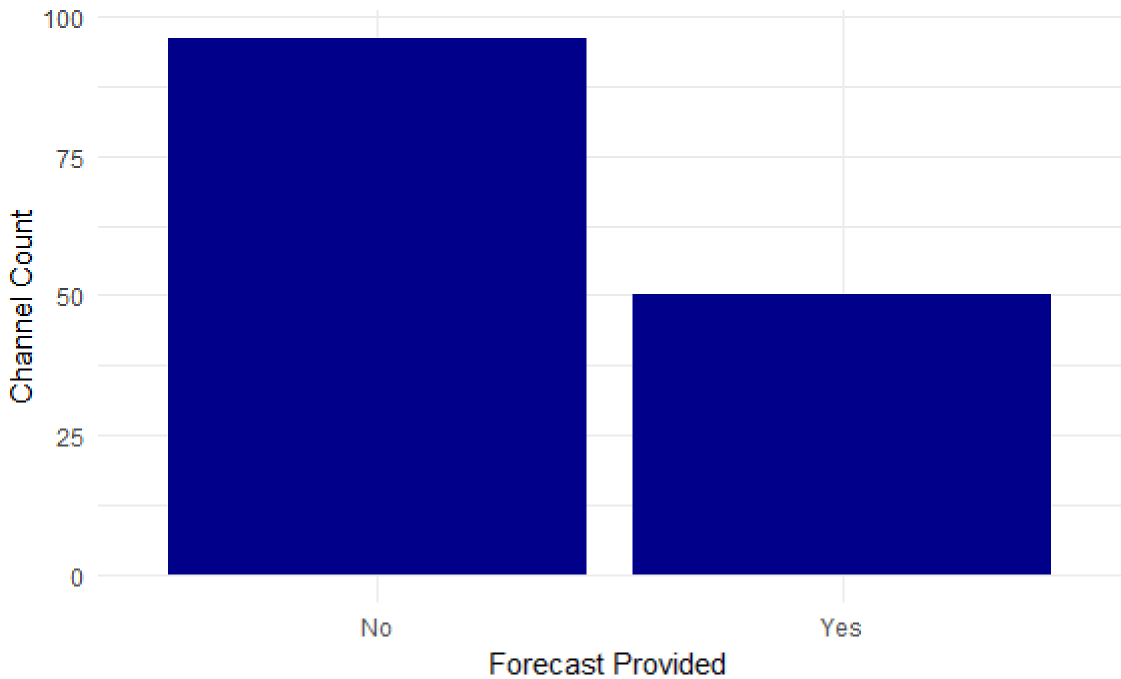


Figure 20 Frequency of days ahead AQ forecasts provided

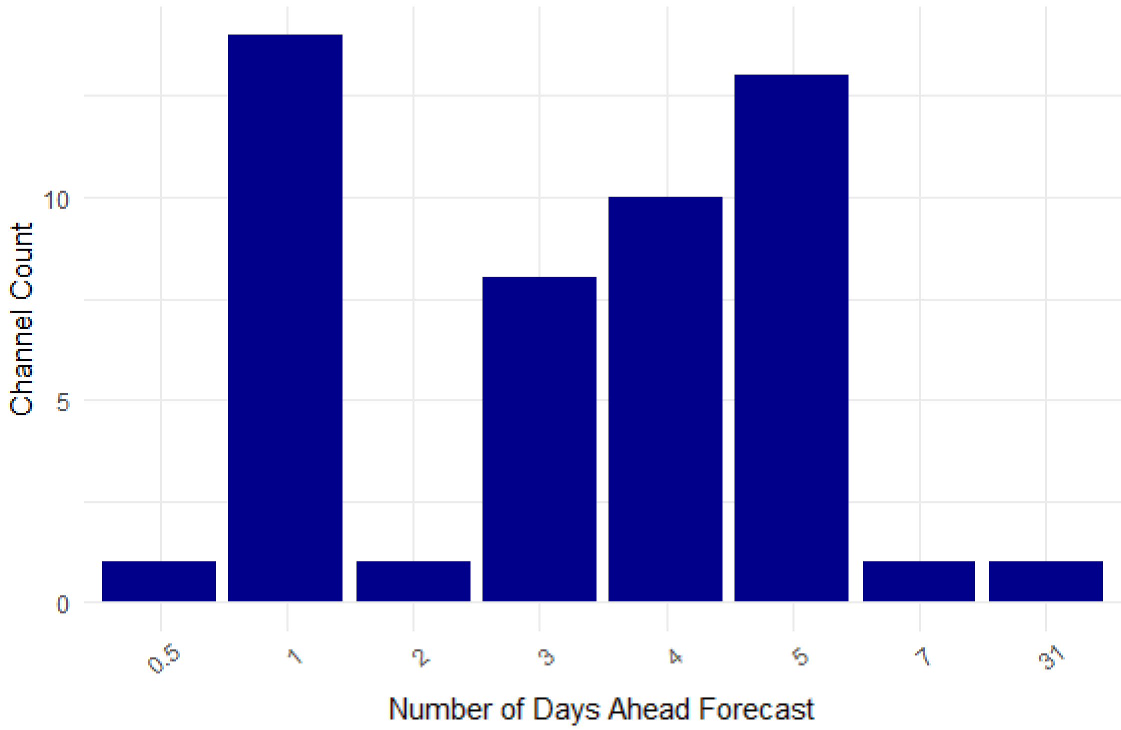
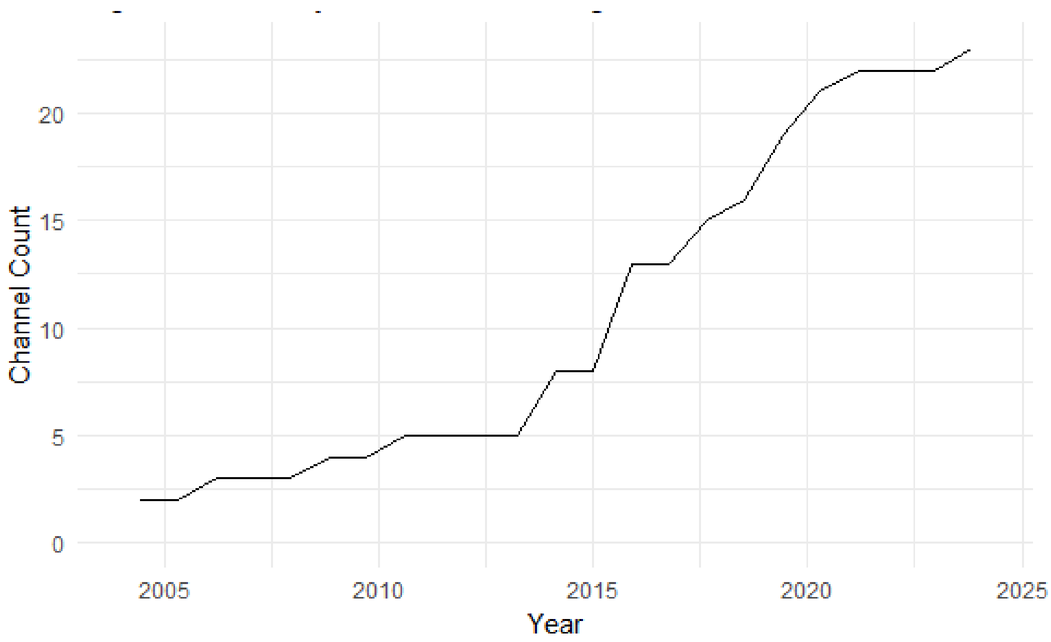


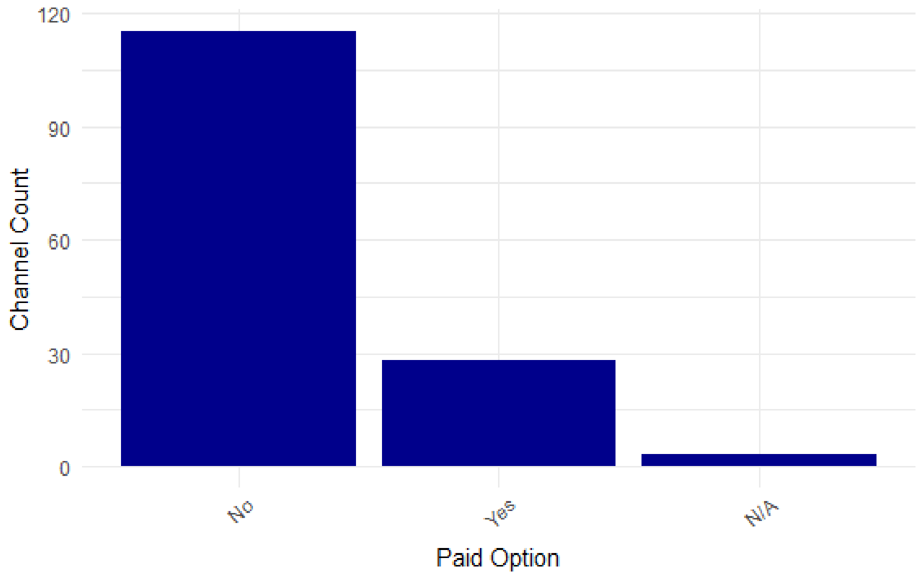
Figure 21 Growth of AQ Channels offering an API over time



Another notable trend is the increase in the number of AQ channels that offer an API (Figure 22). Since 2003, this number has grown from 4 to 26, representing 18% of AQ channels that offer access to an API to obtain AQ data. API access can be offered at no cost or for a cost, most often via a subscription-based model. 19.2% of AQ channels offer an option to pay for premium services or access to their API (Figure 23). Figure

23 does not indicate where AQ channels are advertising consumer products such as air filters, AQ sensors, masks or otherwise.

Figure 22 AQ channels with paid option



5.3 Temporal & geographic resolution of AQ channels

The results demonstrate that AQ channels report on AQ at different spatial and temporal resolutions. The most common time interval for reporting AQ was 'hourly', with over 100 channels delivering information at this rate (Figure 24). Certain channels report air pollution at multiple temporal resolutions. For example, Plume Labs reports near-real time levels as well as a monthly and annual average. Temporal frequency at which AQ is reported is also function of the method used to generate the reading (Figure 26).

AQ channels also provide air pollution data at different geographic resolutions, with the most common resolution being a single latitude/longitude point, followed by a 10 km^2 resolution. Figure 25 shows the geographic resolutions reported across the 146 AQ channels, where certain channels may report air pollution levels at multiple geographic resolutions. 8 AQ channels (5.5%) are dedicated to a specific regional geographic area, for example Sussex, Oxfordshire, and Kent. The majority of the AQ channels returned from the systematic search are available globally. That is, they provide air pollution data for locations around the world. Figure 24 shows the number of channels that are UK-specific (63) compared to those available globally (82).

Figure 23 Temporal resolution of air pollution information reporting across AQ channels

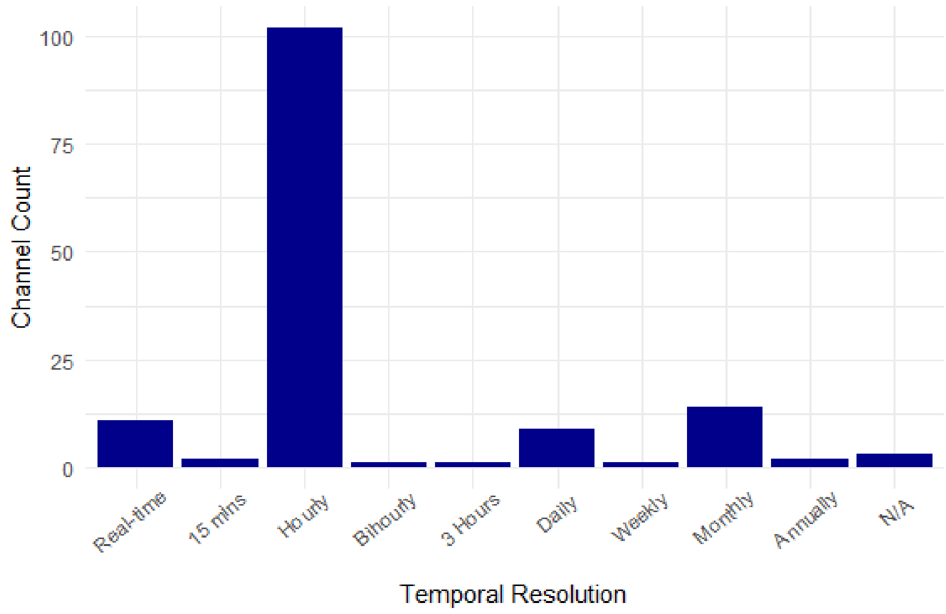


Figure 24 Geographic resolution of air pollution information reporting across AQ channels

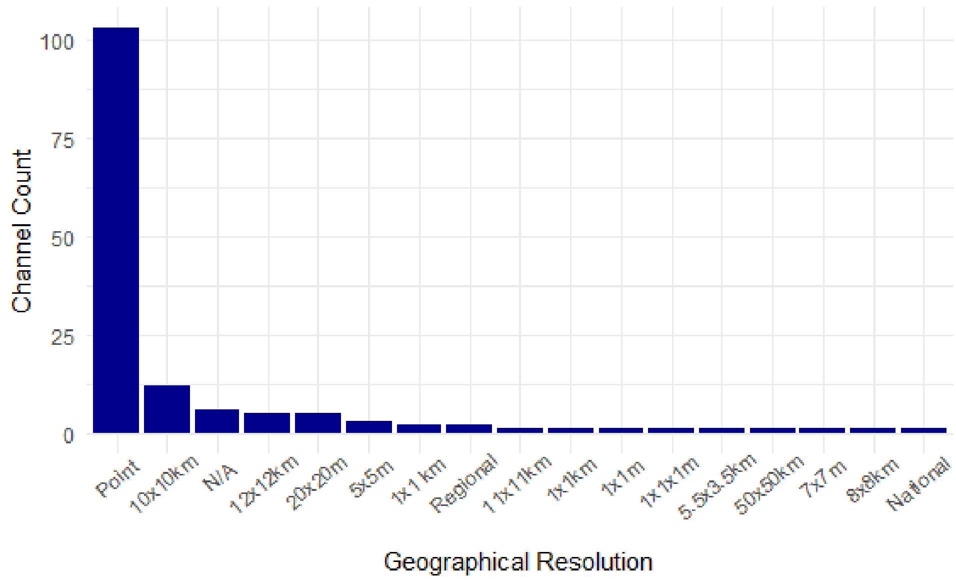


Figure 25 Temporal resolution of data reporting across data source types

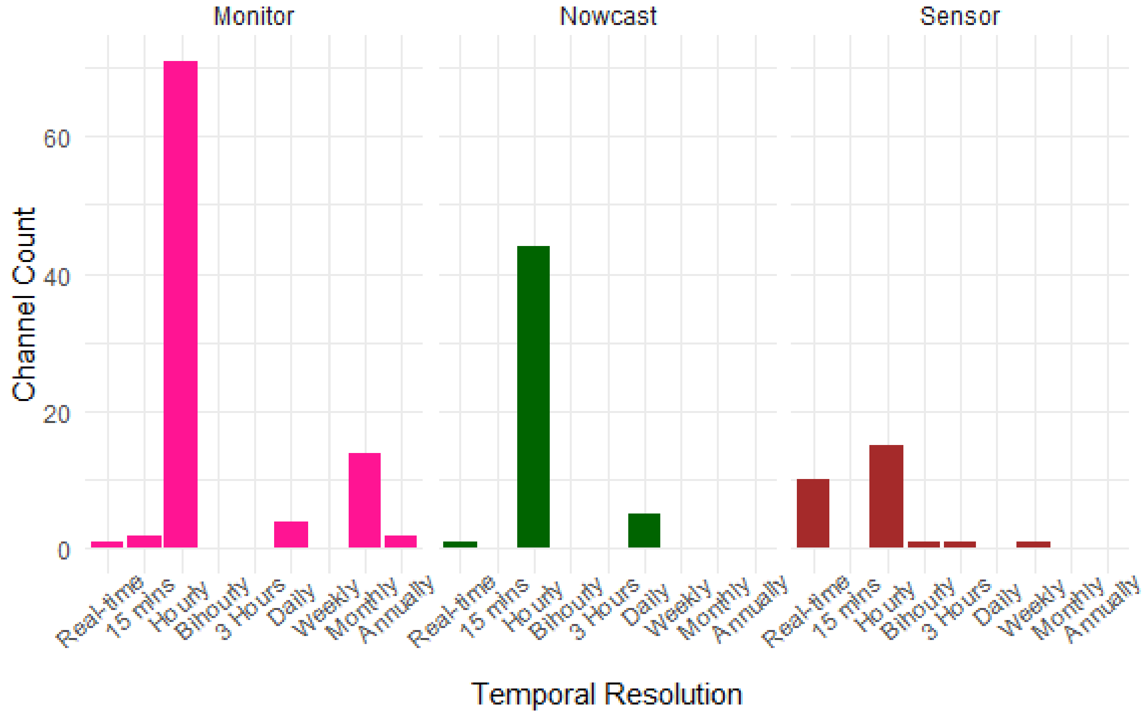
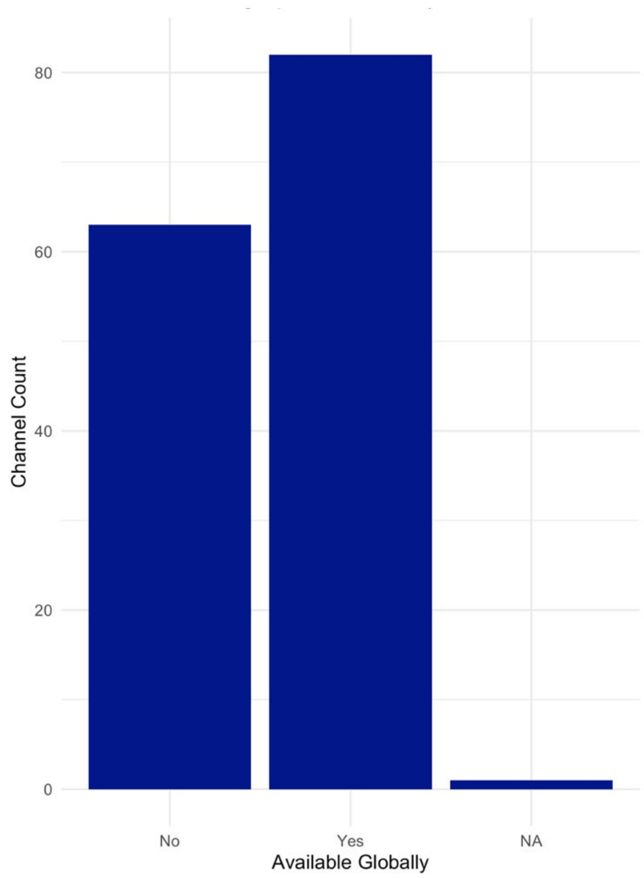


Figure 26 AQ Channel international availability



5.4 Format and characteristics of information presented across AQ channels

Html text data from the homepages of over 90% of channels returned via Google searches, as well as descriptions contained in the app store for app-based channels, were compared with another crosswalk of terms pertaining to key pollutants. The results demonstrated the greatest number of channels included the terms 'AQI', then NO_2, $PM_{2.5}$, followed by PM_{10} (Figure 28).

Figure 27 Pollutant terms reported across AQ channels

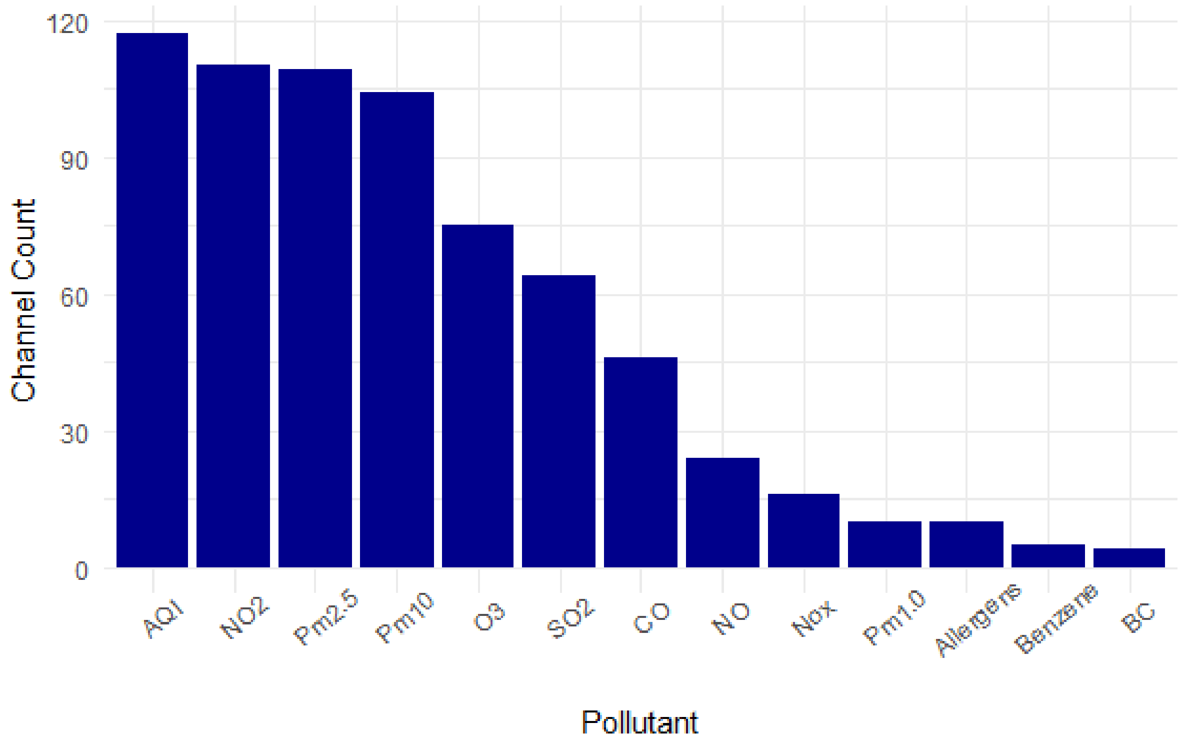


Table 2. Pollution abbreviation key

Abbreviation	Pollutant term
AQI	Air quality index
$PM_{2.5}$	Particulate matter at 2.5 micrograms per cubic meter
PM_{10}	Particulate matter at 10 micrograms per cubic meter
O_3	Ozone
SO_2	Sulphur dioxide
CO	Carbon monoxide
NO	Nitrogen oxide
NOx	Nitrogen oxides
$PM_{1.0}$	Particulate matter at 1 micrograms per cubic meter
BC	Black carbon

AQI is the dominant format for reporting air pollution information, followed by NO2, PM2.5 and PM10 (Figure 28). A subset of the channels provide information about specific pollutants and allergens. Within the AQI subset, the US AQI format was most common AQI adopted, followed by the DAQI and European CAQI (Figure 29).

Figure 28 AQI formats delivered across channels

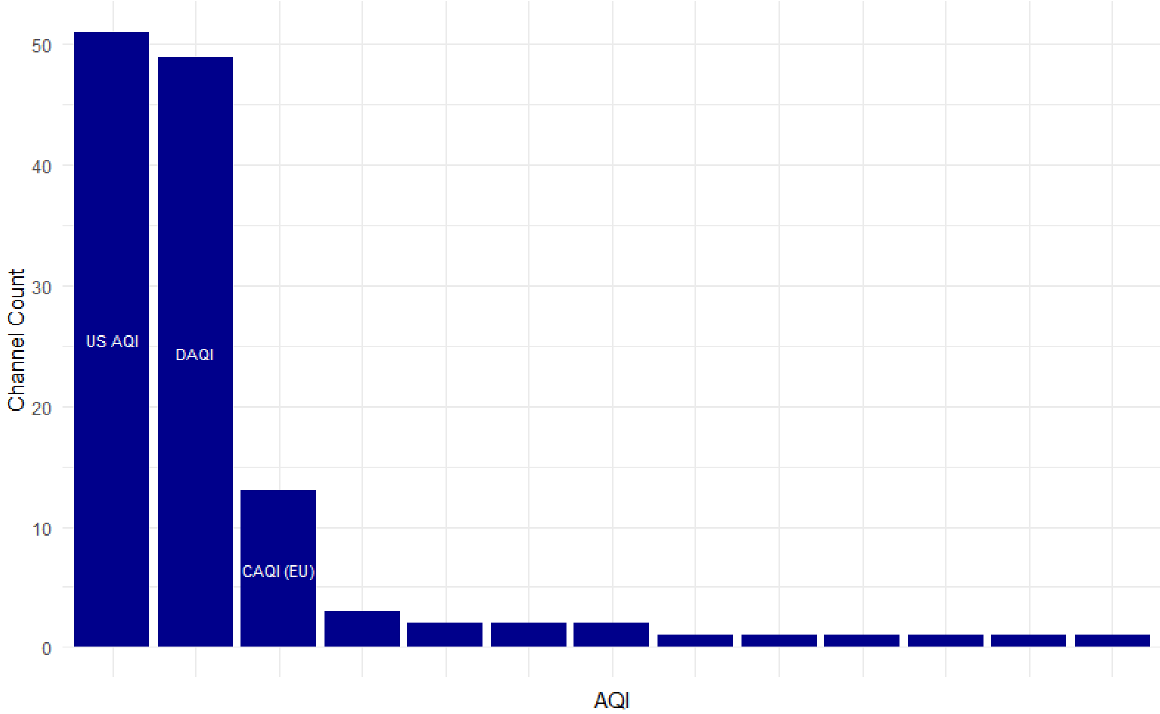
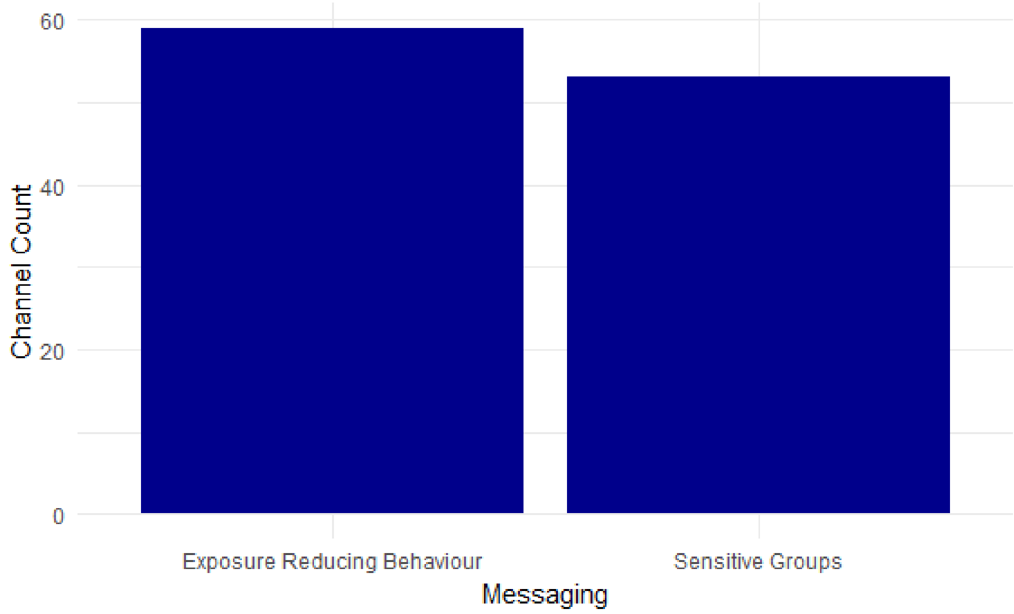


Table 3. AQI abbreviation key

Abbreviation	Description
US AQI	United States Air Quality Index
DAQI	Daily Air Quality Index
CAQI	European Air Quality Index

Figure 29 Number of AQ channels delivering exposure reducing and health-protective messaging



AQIs are linked to specific health messaging, however this context was not necessarily provided through the channel. Figure 30 demonstrates that less than half of the AQ channels highlighted information for sensitive groups, while just under half provided exposure reducing behavioural recommendations.

5.5 AQ channel reach

Metadata pertaining to user access and engagement with AQ channels was limited. As a proxy for evaluating AQ channel reach or popularity, data pertaining to app downloads was obtained when available via the Google Play store. A sample 'Reach dataset' was generated and details are provided in Table 2.

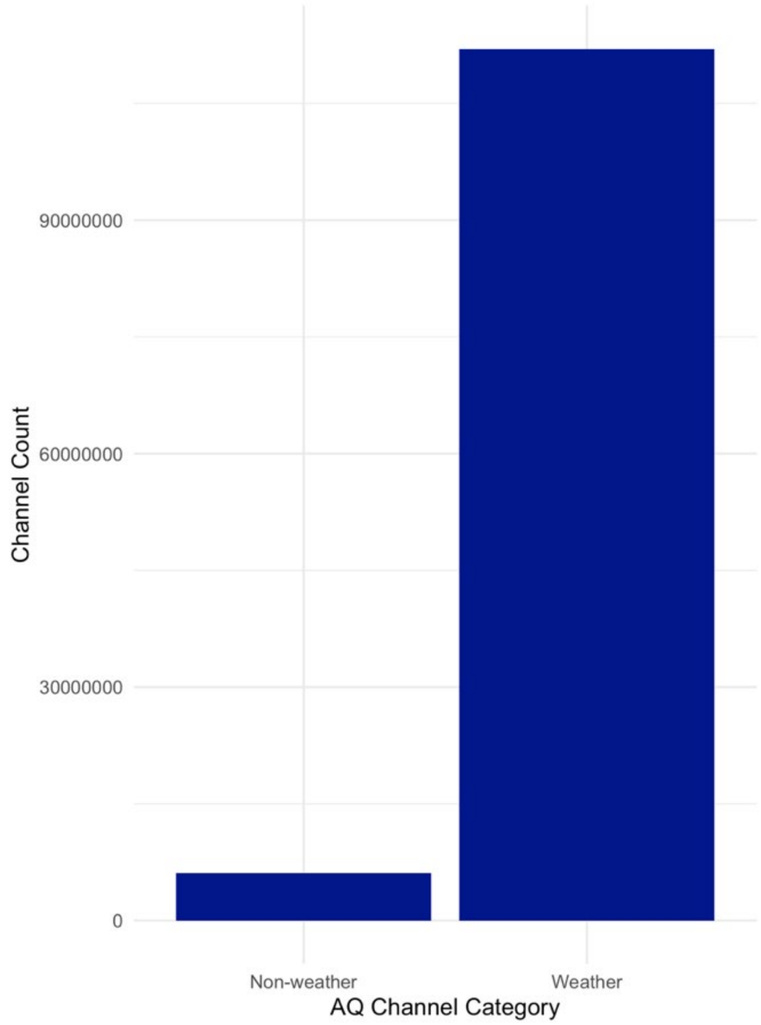
Table 4 Reach dataset summary statistics

	Complete AQ Channel Dataset	**Reach Dataset**
Number of AQ Channels	146	22
Percentage of total dataset	100%	15%
Channel launch date range	2008-12-22 to 2022-08-14	2001-05-09 to 2023-09-20
Percentage available through website platform	88%	55%
Percentage available through phone	63%	100%
Percentage available through tablet	50%	82%
Percentage available through television	1%	5%
Percentage available through smartwatch	14%	32%
Percentage available through Android platform	27%	100%
Percentage available through Apple IoS platform	57%	68%
Percentage offering an API	16%	14%
Percentage using a model data source	34%	27%

Percentage using a monitor data source	64%	59%
Percentage using a sensor data source	19%	32%
Percentage providing a forecast	34%	36%
Percentage providing historic data	51%	32%
Percentage providing PM10 data	71%	82%
Percentage providing PM2.5 data	75%	77%
Percentage providing PM1.0 data	7%	5%
Percentage providing nitrogen dioxide data	75%	68%
Percentage providing sulphur dioxide data	44%	27%
Percentage providing ozone data	51%	41%
Percentage providing carbon monoxide data	32%	18%
Percentage providing nitric oxide data	16%	9%
Percentage providing nitrogen oxides data	11%	0%
Percentage providing black carbon data	3%	0%
Percentage providing benzene data	3%	0%
Percentage providing allergen data	7%	5%
Percentage providing exposure reducing behaviour messaging	40%	68%
Percentage providing sensitive group identification	36%	68%
Percentage providing a paid option	15%	18%

The number of downloads for weather-focused vs. non-weather focused channels was 18 times higher, with a collective 112,000,000 downloads since the channels' inceptions (Figure 31).

Figure 30 Number of downloads for weather-focused vs. non-weather focused channels



5.6 Social media

SerpAPI was used to conduct a search of YouTube using the list of search terms in the Methodology. It is not possible to set a geographic location below the country level when querying YouTube through SerpAPI. At the time the searches were conducted (June 2023), Canadian wildfires were affecting northeastern US which had issued AQ alerts for the region. All results returned for the search strings returned news stories related to the wildfires. YouTube (and other social media platforms) will return different results for the same search term depending on a combination of the user's search history, likes and comments, current news stories and paid promotion. It is not possible therefore to create a reproducible set of results for each search term. The risks and opportunities of social media platforms are explored further in the Discussion.

Twitter's API was not accessible through SerpAPI. Twitter's API changed in March 2023 when API access was stratified into 3 levels of access: Free, Basic ($100/month) and Pro ($5,000/month) (https://developer.twitter.com/en/docs/twitter-api/getting-started/about-twitter-api). Twitter's previous API version (1.1) offered an endpoint that returned users based on a search term (GET users/search). However, this endpoint is not available in the current version (2.0). Instead, researchers used Twitter's advanced search page which returns 20 users based on a search term. Top 20 users for each search term were extracted manually. These were filtered down to 219 unique accounts. These were then assessed manually against the inclusion criteria. Of these accounts, only 3 were classed as AQ channels delivering real-time, forecasted, or historic data at a geographic location. Of these, only 1 was unique to Twitter.

Meta's (Facebook/Instagram) researcher API Fort requires an application process (https://fort.fb.com/researcher-platform). Researchers signed for access in June 2023 but did not hear back in time for this source to be included in this rapid review.

TikTok's research portal was only open to researchers based at academic institutions in the US in June 2023, and therefore was also not included as a source for inclusion in this review.

5.7 Summary

The results of this study show that the digital AQ information landscape and its underlying systems have undergone a significant evolution over the past 15 years marked by several broad trends:

- An increase by a factor of 11 in the number of channels available between 2001 and 2023.
- An increase in the number of organisations providing AQ data over this period, which now stands at 81.
- An increase in the number of private companies providing AQ information.
- A significant increase in the number of channels available as smartphone apps now rivalling websites as the dominant platform.
- An increase in the number of channels providing data from small and/or personal sensors.
  An increase by a factor of 6 in the number of channels offering AQ data via an API.

6 Discussion

6.1 Exploring drivers of the expanding AQ information landscape

The UK digital AQ landscape experienced an increase in the number of channels by a factor of 11, and an increase in the number and types of organisations producing AQ information. Noticeable spikes in growth were observed in 2001, 2012, 2015, 2019 and 2021 (Figure 7). The precise causes of these surges are multifaceted, however it is plausible, for example, that events such as the Dieselgate scandal in 2015 may have sparked heightened public interest in AQ issues. It is also highly plausible that the growth of the AQ channel landscape is linked to the linear growth of the smartphone and mobile application market in the UK over the past decade and a half[19].

A notable trend across the landscape is the increase in popularity of mobile applications compared to traditional websites. The dominance of Apple apps can be attributed to the ease of developing iOS apps, emphasised by the platform's developer-friendly interface. It is possible that the potential for in-app purchases as a revenue stream has led to a proliferation of free apps, contributing to the commercialisation of the AQ information sphere. Companies selling sensors further illustrate this trend, as the monetization potential extends beyond apps to products such as air filters and masks. This commercialisation, and the commercialisation of products marketed to reduce personal pollution exposure, raise concerns about exposure inequality, as individuals with the financial means to invest in these products may experience a different level of protection compared to those who cannot[1,20,].

6.2 Digital infrastructure underpinning the UK digital AQ landscape

The analysis of metadata reveals that AQ channels in the UK primarily obtain data from regulatory monitors, with a growing contribution from sensor-based sources. This is likely a function of the accessibility of monitoring data through open access government channels, which is then operationalised by other AQ channels or data providers. This can lead to a 'Russian doll effect', where peeling back the external layer of an AQ channel reveals the source of the data underpinning the data being reported. There are cases where an additional layer can be peeled back on the data provider, revealing the core data producer. This phenomenon is demonstrated by Figures 12 – 17, which identify the data providers, and in certain instances the data producers that are at the root of the AQ channel. Figure 12 reveals that in most cases, with the exception of PurpleAir, the data producer at the root of AQ channels is a publicly funded, multi-institutional venture.

The 'reach' sample dataset illustrates how AQ information carried in weather-focussed apps and websites dominates the landscape. Weather-focussed apps have been downloaded over 100 million times versus over 6 million times for AQ-specific apps downloaded. It is worth noting that IQAir makes up 5 million of the AQ-specific apps. In a landscape where AQ channels vie for public attention, the asymmetry between the resources available to government-backed entities versus corporate counterparts becomes apparent.

The UK's digital AQ information landscape reveals a surge in the availability of APIs for accessing AQ information. This increase indicates a growing interest in providing real-time, customizable data feeds to users. Simultaneously, the influx of commercial entities into the space, launching both apps and websites focused on AQ, is linked to the rise of personal sensors. These channels often serve as marketing tools for companies offering supposed pollution-abating products. AQ channels, in this context, act as a conduit for promoting products such as air filters and masks, or even sensor data for personal exposure reduction.

6.3 Characteristics of information & messaging

The most common time interval for reporting AQ is 'hourly'. This aligns with the standard of hourly reporting established by Defra in 1987[21,22]. The most common resolution being a single latitude/longitude point, followed by a 10km^2 resolution. An AQ channel may report pollution at multiple geographic resolutions. This is most often the case when a channel derives data from both sensors or monitors (point) and a model (varying spatial resolutions).

The majority of AQ channels deliver readings for a specific point location. This can contribute to confusion around how to interpret an air pollution reading. For example, if an individual has their location services[iii] activated, one may assume that the reading delivered by an AQ channel corresponds to their exact location, where in fact the channel is reporting data from the nearest monitor or sensor, or 1km or greater resolution model.

The predominant use of the US Air Quality Index (AQI) format across various channels is a notable feature of the digital AQ information landscape (Figure 29). This adoption is likely due to the global prevalence of the US AQI, making it a default choice for many platforms. However, a diverse range of indices is also observed, introduced by specific channels. This diversity contributes to a richer information landscape but also raises questions about standardisation and the potential for confusion among users.

The global nature of the digital society we inhabit is reflected in the majority of AQ channels being accessible worldwide (Figure 27). Many channels were not UK-specific and are likely developed outside of the UK presenting a greater challenge in terms of adoption of UK-specific indices and related messages.

6.4 Emerging trends

The trends revealed in this review reveal a multifaceted and evolving AQ information landscape. The provision of APIs and subscription-based services signifies a shift towards more interactive and personalized data access.

Social media has become a pivotal platform for information dissemination in the UK. Research from Kantar reveals that 63% of the UK population accesses social media every day, with users spending around 20 minutes a day across Facebook, Twitter/X and Instagram and around 30 minutes per day dedicated to YouTube. 47% of UK adults say they find out about breaking news through social media but only 30% of people trust social media as a news source[23].

Influencers on Twitter/X, Instagram, TikTok and YouTube have millions of followers and represent another opportunity for AQ information dissemination. Existing research signals the potential for social media to affect health outcomes[24]. These platforms are video-centric, emphasising the need for communication methodologies that can handle visual and video content. The information landscape on social media is nuanced however, as algorithms mediate user interactions, making it essential to distinguish between direct and algorithmically mediated channels. While our results indicate a limited number of channels offering AQ information on social media, this clearly represents an untapped opportunity.

Real-time data is gaining emphasis, aligning with the 'smart' and Internet of Things trends. Communities are increasingly participating in monitoring AQ themselves,

[iii] Location services refer to a mobile or digital device feature that utilizes a combination of technologies, such as GPS (Global Positioning System), Wi-Fi, cellular networks, and Bluetooth, to determine and track the geographical position of the device. This function enables applications and services to access the device's current location, facilitating location-based features, such as mapping, navigation, geotagging, and location-based notifications.

expanding diversity and engagement, and challenging the traditional role of the government as the sole provider of AQ information. Some AQ channels report 'real time data' (i.e. 1-5 min averages) using the US AQI or DAQI banding system despite the averaging periods being markedly different.

Overall, the landscape of digital AQ information is evolving towards greater interactivity and community engagement and is featuring real or near-real time data at increasingly fine geographic resolutions. The results show that government is no longer the exclusive player in the AQ information space, and understanding user trust and the incorporation of WHO guidelines as the gold standard will be crucial for navigating this dynamic terrain.

6.5 Limitations

The limitations of this digital channels review encompass both the platforms considered and the broader challenges in disseminating AQ information. While WhatsApp has emerged as a crucial communication tool, especially during COVID with the introduction of broadcast systems, its potential as an information source necessitates deeper analysis, particularly from a social networks perspective.

As discussed, YouTube, although initially considered, was excluded from the investigation due to the U.S.-centric content returned in preliminary searches. The exclusion of YouTube highlights the challenges in obtaining diverse and region-specific AQ information through certain platforms.

The inclusion of Twitter data in this review exemplifies the evolving landscape of social media platforms. It also underscores the impact of algorithms and human interactions on the fluidity of inputs and outputs. It's important to note that the Twitter data presented is based on what was accessible without resorting to the expensive professional API, acknowledging the potential limitations in the comprehensiveness of the data.

Due to rapid nature of this study, this review was not able to investigate the prevalence of AQ information carried on broadcast radio and TV. Despite recent declines, these remain important sources of information in the UK reaching 80-90% of the UK population each week[25]. The engagement of broadcast TV and radio with AQ information should be considered more fully in future work.

Additionally, outdoor displays such as the GLA emergency alert system and campaigns delivered via dynamic, data-linked billboards[14,16] has emerged as an AQ communication trend. Identifying and examining these case examples requires a more specific methodology that can account for visual media.

This review explicitly focuses on digital channels, however digital exclusion poses a significant challenge for some key demographics in the UK. Disseminating AQ information solely through digital channels may not be sufficient, particularly for individuals who do not have internet access or those do not engage with digital technologies (i.e. older individuals). AQ information communication strategies must be expanded to involve social actors such as general practitioners, schools, and workplaces to ensure maximum reach to diverse groups throughout the UK. This acknowledgment underscores the multifaceted nature of information dissemination and emphasises the need for a comprehensive approach that integrates digital platforms with traditional channels to address the complexities of reaching varied demographics.

7 Conclusion

The current digital AQ information landscape in the UK reflects a dynamic interplay between government-funded data production and commercial AQ channels. The increase in channels observed from 2008 to 2023, particularly in mobile applications, is likely driven by commercialisation trends including in-app purchases and air pollution sensor sales. This is also mirrored in widespread API availability through subscription-based services marketed to both individuals and businesses. These features raise important questions around equity of information access, the standardisation of indices, and the interconnectedness of environmental issues in our digital society.

Several key recommendations emerged from this rapid evidence review. First, leveraging existing channels rather than creating new ones can optimize resources and streamline information dissemination. Additionally, the potential of social media as a powerful communication tool should be carefully considered, recognising its personalised nature and the need for targeted outreach.

Moreover, recognising the contemporary trend of personalised information delivery, Defra should consider establishing a visual media presence, potentially through videos, to engage a wider audience. Additionally, the adoption of Application Programming Interfaces (APIs) by leading data providers suggests a potential avenue for Defra to enhance the reach of its AQ index and messaging. It is worth considering how developments at the backend, i.e. through provision of an API for seamless integration into existing platforms, can amplify the impact and clarity of AQ information. Conditions around government API usage, particularly around generating AQ messaging from queried data or leveraging of publicly funded data to support the marketing of commercial products should be evaluated further.

Given the evolving diversity within the AQ landscape since the last update of the Daily Air Quality Index (DAQI), Defra should also consider implementing an engagement strategy with key providers in the space. This may include collaborating with channels that are weather-focused, which demonstrated the greatest number of users. Adopting a collaborative approach could help support the rollout of an updated DAQI, ensuring that the information remains relevant and accessible to the public.

In addressing the increasing commercialisation of AQ channels, it is crucial to navigate the balance between public interest and private enterprise. The review highlights the need for careful consideration and collaborative efforts to uphold the integrity of AQ information in the digital age. Ultimately, these recommendations aim to fortify government's role in shaping and disseminating accurate and accessible AQ information while navigating the complexities of a rapidly evolving digital landscape.

7.1 Summary

- The number of AQ channels has increased substantially since 2001 (by a factor of 11X)
- The increase in AQ channels is likely driven by technological and data sharing innovations reflected in the number of data producers and providers offering APIs and commercialisation of personal exposure interventions.
- The increase in AQ channels over time may also be influenced or coproduced alongside greater attention to air pollution throughout national and international media.
- Leveraging existing channels rather than creating new ones can optimize resources and streamline information dissemination.
- There is a need to advance methods for identifying cases, characteristics, and effects of AQ information sharing across social media platforms, radio, TV, and public outdoor displays.
- There is a need to advance methods for examining the social, interpersonal and cultural dimensions of information sharing about AQ.
- Further work is warranted to examine qualities of AQ channels delivered by AQ channels coming from government, NGO, or commercial entities.

7.2 Future research

Future research should build on findings from this work and explore the prevalence of the most recent World Health Organization (WHO) guidelines (2021) across UK AQ channels, while considering its potential as a reporting metric for AQ in the UK. Evaluations of this could involve research of user experiences with initiatives like AddressPollution, which provides long-term, postcode-level data and features WHO guidelines. showcases a move towards using the WHO guidelines as the standard reporting metric for AQ in the UK.

Additional research is warranted around how the presentation of AQ information differs across government, NGOs, and commercial entities. In addition, trust in information sources is a critical aspect, and further research is needed to understand how users perceive and trust channels from different categories of providers, including government, business, and healthcare providers.

Future research should explore the prevalence and impact of AQ information dissemination through broadcast radio and TV, considering their continued significance in reaching a large portion of the UK population. Additionally, digital billboards and visual displays should be evaluated as mechanisms for conveying AQ information. Developing specific methodologies to identify, analyze, and assess the impact of visual displays can enhance our understanding of their role in conveying real-time air quality updates and promoting public responsiveness. Future studies should explore effective methodologies for evaluating the visual elements of AQ communication, considering factors such as design, placement, and message framing to provide practical insights for optimizing outdoor displays in air quality awareness initiatives.

A more comprehensive evaluation of social media is also required. This should investigate the role of social networks, social dynamics and informational credibility throughout digital space. Future work in this area can provide important insight around how certain social media platforms are already involved in the communication of AQ information, and could support addition communication ventures.

8 Acknowledgements

We are grateful to the air quality team at Defra for commissioning this review. Thank you to our internal reviewers Dr Mohammed Mead, Tim Baker, Professor Ben Barratt and Professor Frank Kelly. Thank you also to Valentina Lotti for her help recruiting research assistants Adam Brighty and William Francis.

9 Appendix

9.1 Search terms

Search terms for Google search, Apple IOS, and Google App stores:

- Air quality
- Air quality app
- Air quality forecast
- Air quality nowcast
- Air quality alert
- Air quality map
- Air pollution
- Air pollution app
- Air pollution forecast
- Air pollution nowcast
- Air pollution alert
- Air pollution map
- Smog
- Smog app
- Smog air
- Smog alert
- Smog forecast
- Smog nowcast
- Check air quality
- Air quality today
- Air quality near me
- My AQI air

9.2 Search locations

Table 3 below indicates the location name taken from the supported location list. The cells highlighted green indicate which name for each location was able to be read by SerpAPI.

Table 5 Geographic search locations

Location	County/Largest Town	Postcode
Bath,England,United+Kingdom		
Birmingham,England,United+Kingdom		
Bradford,England,United+Kingdom		
Brighton+&+Hove,England,United+Kingdom		
Bristol,England,United+Kingdom		
Cambridge,England,United+Kingdom		
Canterbury,England,United+Kingdom		
Carlisle,England,United+Kingdom		
Chelmsford,England,United+Kingdom		
Chester,England,United+Kingdom		
Chichester,England,United+Kingdom		
Colchester,England,United+Kingdom		
Coventry,England,United+Kingdom		
Derby,England,United+Kingdom		
Doncaster,England,United+Kingdom		
Durham,England,United+Kingdom		
Ely,England,United+Kingdom		
Exeter,England,United+Kingdom		
Gloucester,England,United+Kingdom		
Hereford,England,United+Kingdom		
Kingston-upon-Hull,England,United+Kingdom		
Lancaster,England,United+Kingdom		
Leeds,England,United+Kingdom		
Leicester,England,United+Kingdom		
Lichfield,England,United+Kingdom		
Lincoln,England,United+Kingdom		
Liverpool,England,United+Kingdom		
London,England,United+Kingdom		
Manchester,England,United+Kingdom		
Milton+Keynes,England,United+Kingdom		
Newcastle-upon-Tyne,England,United+Kingdom		
Norwich,England,United+Kingdom		
Nottingham,England,United+Kingdom		
Oxford,England,United+Kingdom		

Peterborough,England,United+Kingdom		
Plymouth,England,United+Kingdom		
Portsmouth,England,United+Kingdom		
Preston,England,United+Kingdom		
Ripon,England,United+Kingdom		
Salford,England,United+Kingdom		
Salisbury,England,United+Kingdom		
Sheffield,England,United+Kingdom		
Southampton,England,United+Kingdom		
Southend-on-Sea,England,United+Kingdom		
St+Albans,England,United+Kingdom		
Stoke+on+Trent,England,United+Kingdom		
Sunderland,England,United+Kingdom		
Truro,England,United+Kingdom		
Wakefield,England,United+Kingdom		
Wells,England,United+Kingdom		
Winchester,England,United+Kingdom		
Wolverhampton,England,United+Kingdom		
Worcester,England,United+Kingdom		
York,England,United+Kingdom		
Armagh,Northern+Ireland,United+Kingdom		
Bangor,Northern+Ireland,United+Kingdom		
Belfast,Northern+Ireland,United+Kingdom		
Lisburn,Northern+Ireland,United+Kingdom		
Londonderry,Northern+Ireland,United+Kingdom		
Newry,Northern+Ireland,United+Kingdom		
Aberdeen,Scotland,United+Kingdom		
Dundee,Scotland,United+Kingdom		
Dunfermline,Scotland,United+Kingdom		
Edinburgh,Scotland,United+Kingdom		
Glasgow,Scotland,United+Kingdom		
Inverness,Scotland,United+Kingdom		
Perth,Scotland,United+Kingdom		
Stirling,Scotland,United+Kingdom		
Bangor,Wales,United+Kingdom		
Cardiff,Wales,United+Kingdom		
Newport,Wales,United+Kingdom		
St+Asaph,Wales,United+Kingdom		
St+Davids,Wales,United+Kingdom		
Swansea,Wales,United+Kingdom		
Wrexham,Wales,United+Kingdom		
Barking+and+Dagenham,London,England,United+Kingdom		IG11

Barnet,London,England,United+Kingdom		EN4
Bexley,London,England,United+Kingdom		DA14
Brent,London,England,United+Kingdom		HA0
Bromley,London,England,United+Kingdom		BR1
Camden,London,England,United+Kingdom		NW1
City+of+London,London,England,United+Kingdom		
Croydon,England,United+Kingdom		BR3
Ealing,London,England,United+Kingdom		W5
Greenwich,London,England,United+Kingdom		E14
Hackney,London,England,United+Kingdom		E1
Hammersmith+and+Fulham,England,United+Kingdom		SW6
Haringey,London,England,United+Kingdom		N10
Harrow,London,England,United+Kingdom		HA1
Havering,London,England,United+Kingdom		RM1
Hillingdon,London,England,United+Kingdom		HA2
Hounslow,London,England,United+Kingdom		TW4
Islington,London,England,United+Kingdom		N1
Royal+Borough+of+Kensington+and+Chelsea,England,United+Kingdom		SW10
Kingston+upon+Thames,England,United+Kingdom		KT1
Lambeth,London,England,United+Kingdom		
Lewisham,London,England,United+Kingdom		SE13
Merton,London,England,United+Kingdom		CR4
Newham,London,England,United+Kingdom		E12
Redbridge,London,England,United+Kingdom		E11
Richmond+upon+Thames,England,United+Kingdom		KT2
Southwark,London,England,United+Kingdom		SE1
Sutton,England,United+Kingdom		SM1
Tower+Hamlets,London,England,United+Kingdom		E1
Waltham+Forest,London,England,United+Kingdom		E10
Wandsworth,London,England,United+Kingdom		SW11
Westminster,England,United+Kingdom		SW1A
Bedfordshire,England,United+Kingdom		
Berkshire,England,United+Kingdom		
Bristol,England,United+Kingdom		
Buckinghamshire,England,United+Kingdom		
Cambridgeshire,England,United+Kingdom		
Cheshire,England,United+Kingdom		
Cornwall,England,United+Kingdom		
Cumbria,England,United+Kingdom		
Derbyshire,England,United+Kingdom		
Devon,England,United+Kingdom		
Dorset,England,United+Kingdom		

County+Durham,England,United+Kingdom		
East+Riding+of+Yorkshire,England,United+Kingdom		
East+Sussex,England,United+Kingdom		
Essex,England,United+Kingdom		
Gloucestershire,England,United+Kingdom		
Greater+London,England,United+Kingdom		
Greater+Manchester,England,United+Kingdom		
Hampshire,England,United+Kingdom		
Herefordshire,England,United+Kingdom		
Hertfordshire,England,United+Kingdom		
Isle+of+Wight,England,United+Kingdom		
Kent,England,United+Kingdom		
Lancashire,England,United+Kingdom		
Leicestershire,England,United+Kingdom		
Lincolnshire,England,United+Kingdom		
City+of+London,England,United+Kingdom		
Merseyside,England,United+Kingdom		
Norfolk,England,United+Kingdom		
Northamptonshire,England,United+Kingdom		
Northumberland,England,United+Kingdom		
North+Yorkshire,England,United+Kingdom		
Nottinghamshire,England,United+Kingdom		
Oxfordshire,England,United+Kingdom		
Rutland,England,United+Kingdom		
Shropshire,England,United+Kingdom		
Somerset,England,United+Kingdom		
South+Yorkshire,England,United+Kingdom		
Staffordshire,England,United+Kingdom		
Suffolk,England,United+Kingdom		
Surrey,England,United+Kingdom		
Tyne+and+Wear,England,United+Kingdom		
Warwickshire,England,United+Kingdom		
West+Midlands,England,United+Kingdom		
West+Sussex,England,United+Kingdom		
West+Yorkshire,England,United+Kingdom		
Wiltshire,England,United+Kingdom		
Worcestershire,England,United+Kingdom		
Antrim,Northern+Ireland,United+Kingdom		
Armagh,Northern+Ireland,United+Kingdom	Keady	BT60
Down,Northern+Ireland,United+Kingdom	Downpatrick	
Fermanagh,Northern+Ireland,United+Kingdom	Lisnaskea	BT92
Londonderry,Northern+Ireland,United+Kingdom		

Tyrone,Northern+Ireland,United+Kingdom	Omagh	
Angus,Scotland,United+Kingdom	Forfar	
Argyll,Scotland,United+Kingdom	Inveraray	
Ayrshire,Scotland,United+Kingdom	Ayr	
Banffshire,Scotland,United+Kingdom	Banff	
Berwickshire,Scotland,United+Kingdom	Berwick	
Bute,Scotland,United+Kingdom	Rothesay	
Caithness,Scotland,United+Kingdom	Wick	
Clackmannanshire,Scotland,United+Kingdom	Alloa	
Dumfriesshire,Scotland,United+Kingdom	Dumfries	
Dunbartonshire,Scotland,United+Kingdom	Dumbarton	
East+Lothian,Scotland,United+Kingdom	Haddington	EH41
Fife,Scotland,United+Kingdom		
Inverness-shire,Scotland,United+Kingdom	Inverness	
Kincardineshire,Scotland,United+Kingdom	Stonehaven	AB39
Kinross-shire,Scotland,United+Kingdom	Kinross	
Kirkcudbrightshire,Scotland,United+Kingdom	Kirkcudbright	DG7
Lanarkshire,Scotland,United+Kingdom	Lanark	ML11
Midlothian,Scotland,United+Kingdom	Dalkeith	EH22
Moray,Scotland,United+Kingdom	Elgin	
Nairnshire,Scotland,United+Kingdom	Nairn	
Orkney,Scotland,United+Kingdom	Kirkwall	
Peeblesshire,Scotland,United+Kingdom	Peebles	
Perthshire,Scotland,United+Kingdom	Perth	
Renfrewshire,Scotland,United+Kingdom	Renfrew	
Ross+and+Cromarty,Scotland,United+Kingdom	Dingwall	
Roxburghshire,Scotland,United+Kingdom	Kelso	
Selkirkshire,Scotland,United+Kingdom	Selkirk	TD7
Shetland,Scotland,United+Kingdom	Lerwick	
Stirlingshire,Scotland,United+Kingdom	Stirling	
Sutherland,Scotland,United+Kingdom	Dornoch	
West+Lothian,Scotland,United+Kingdom		
Wigtownshire,Scotland,United+Kingdom	Wigtown	DG8
Blaenau+Gwent,Wales,United+Kingdom		
Bridgend,Wales,United+Kingdom		
Caerphilly,Wales,United+Kingdom		
Camarthenshire,Wales,United+Kingdom	Camarthen	SA31
Ceredigion,Wales,United+Kingdom	Cardigan	SA43
Conwy,Wales,United+Kingdom		
Denbighshire,Wales,United+Kingdom	Denbigh	
Flintshire,Wales,United+Kingdom	Flint	
Gwynedd,Wales,United+Kingdom	Bangor	

Isle+of+Anglesey,Wales,United+Kingdom	Holyhead	LL65
Merthyr+Tydfil,Wales,United+Kingdom		
Monmouthshire,Wales,United+Kingdom	Monmouth	
Neath+Port+Talbot,Wales,United+Kingdom	Neath	
Pembrokeshire,Wales,United+Kingdom	Haverfordwest	
Powys,Wales,United+Kingdom	Newtown	
Rhondda+Cynon+Taf,Wales,United+Kingdom	Aberdare	CF44
Torfaen,Wales,United+Kingdom	Cwmbran	
Vale+of+Glamorgan,Wales,United+Kingdom	Barry	

9.3 Metadata list

Metadata was gathered for the following fields (where available):

- Channel name
- Channel provider
- URL
- Channel launch date
- Delivery mechanism (website, mobile app, radio, television)
- What hardware is it available on (computer, phone, tablet, television, smartwatch)
- If mobile app, which provider (Android, iOS, or both)
- Does the channel offer an API?
- Can it be paired with a personal sensor?
- AQ data source (model, monitor or sensor)
- If model, what type of model (Proprietary or existing government model)
- Model name (if available)
- Geographic resolution of the estimate or measurement
- Temporal frequency of measurements
- Temporal frequency of model updates
- AQ data provider (if different from channel provider)
- Is a forecast available?
- How many days ahead forecasted?
- Is historic data available?
- What pollutants are reported
- Is an AQI shown?
- If an AQI is shown, which one(s)?
- Latest update
- Returned data or reading in the past year?
- Volume of access
- Total downloads (via Google store)
- Total visits in past year (Google analytics/social media views and followers)
- Region (UK or globally available?)
- Are exposure reducing behavioural (ERB) messaging or recommendations provided?
- What messages are provided?

- Are sensitive groups identified?
- How are sensitive groups defined
- Are push notifications available?
- Paid option?

9.3.1 App store metadata

Table 4 below shows which metadata information was provided across Google Play and Apple iOS app stores.

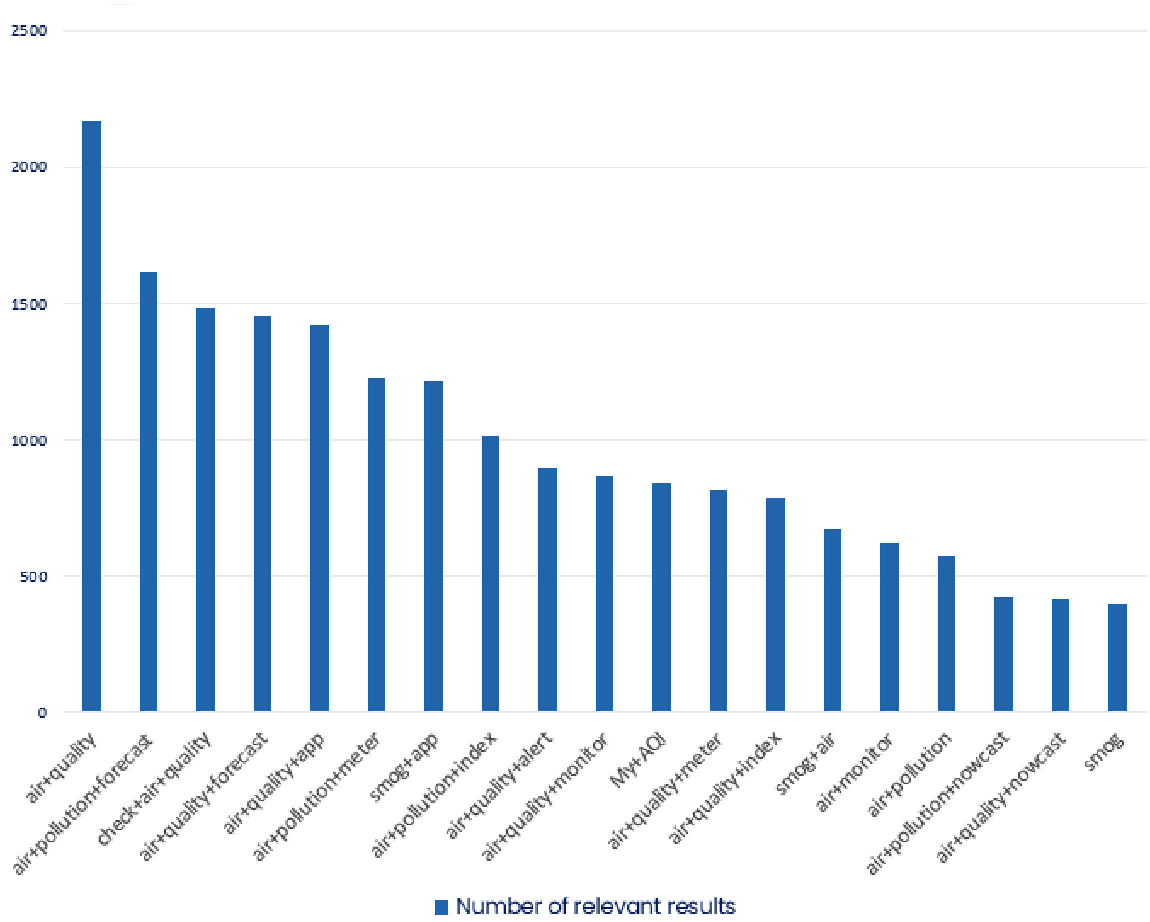
Table 6. App store metadata

App Store	Title	Link	Product ID	Description	Category	Release Date	Price	Downloads
Apple								
Google Play								

9.4 Websites (Google Search results)

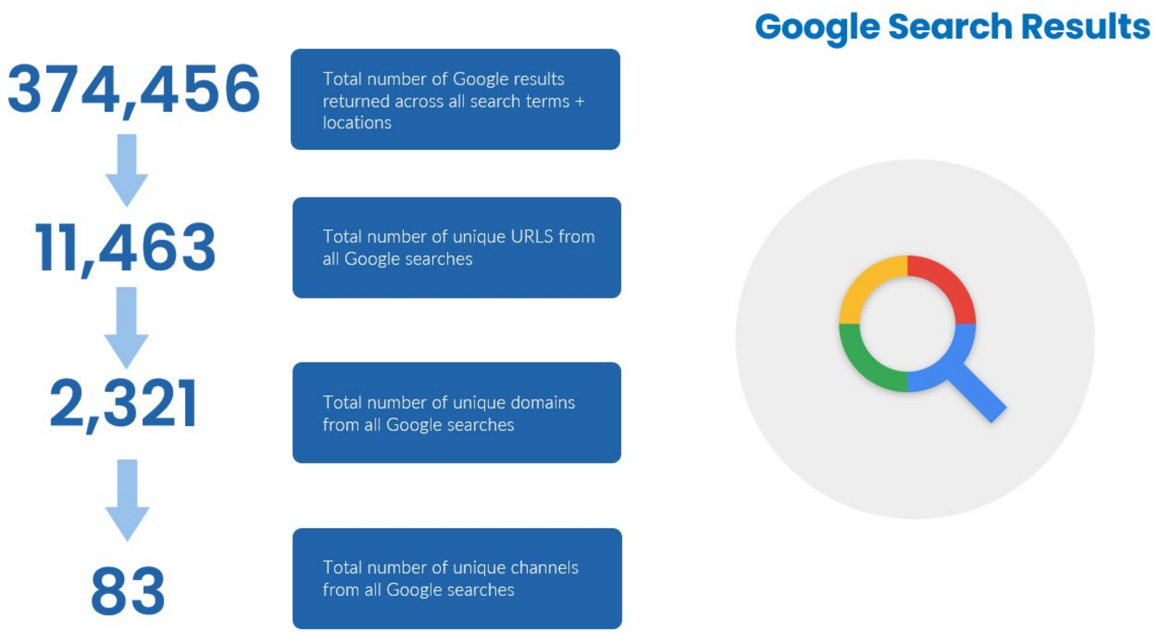
The list of key search terms each returned a different number of relevant results. The strings ‘air pollution forecast’, ‘air quality forecast’, ‘air quality’, ‘air quality app’ and ‘check air quality’ returned the highest number of relevant results.

Figure 31 Relevant results returned from Google search for each key search term



Searches of the list of key terms conducted via the Google search engine across 207 geolocations resulted in 374,456 results (Figure 32). The searches were refined according to the criteria described in the Methodology, section 4.2.1.

Figure 32 Refinement process for results returned via the Google search engine



After the total results were filtered for unique URLs and domains (tier one), the crosswalk of common terms was used to generate a score for each unique result. This score was used filter results were most likely to be AQ channels. This aided the manual refining process, where Google Books links, news domains and clear outliers were filtered out.

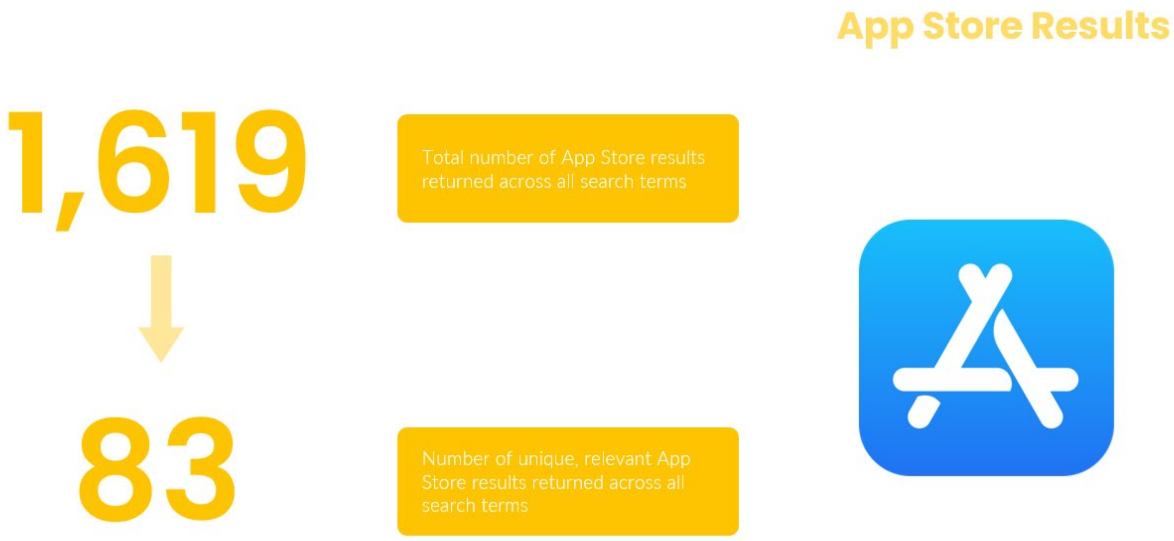
9.5 Smartphone Apps

The app store search results were refined from a total of 1913 total apps returned across the Google Play and Apple iOS app store. The crosswalk of common terms was used to refine the list of apps, resulting in a total of 185. Characteristics of the app-based AQ channels resulting from the refinement process for each store are described separately.

9.5.1 Apple App Store

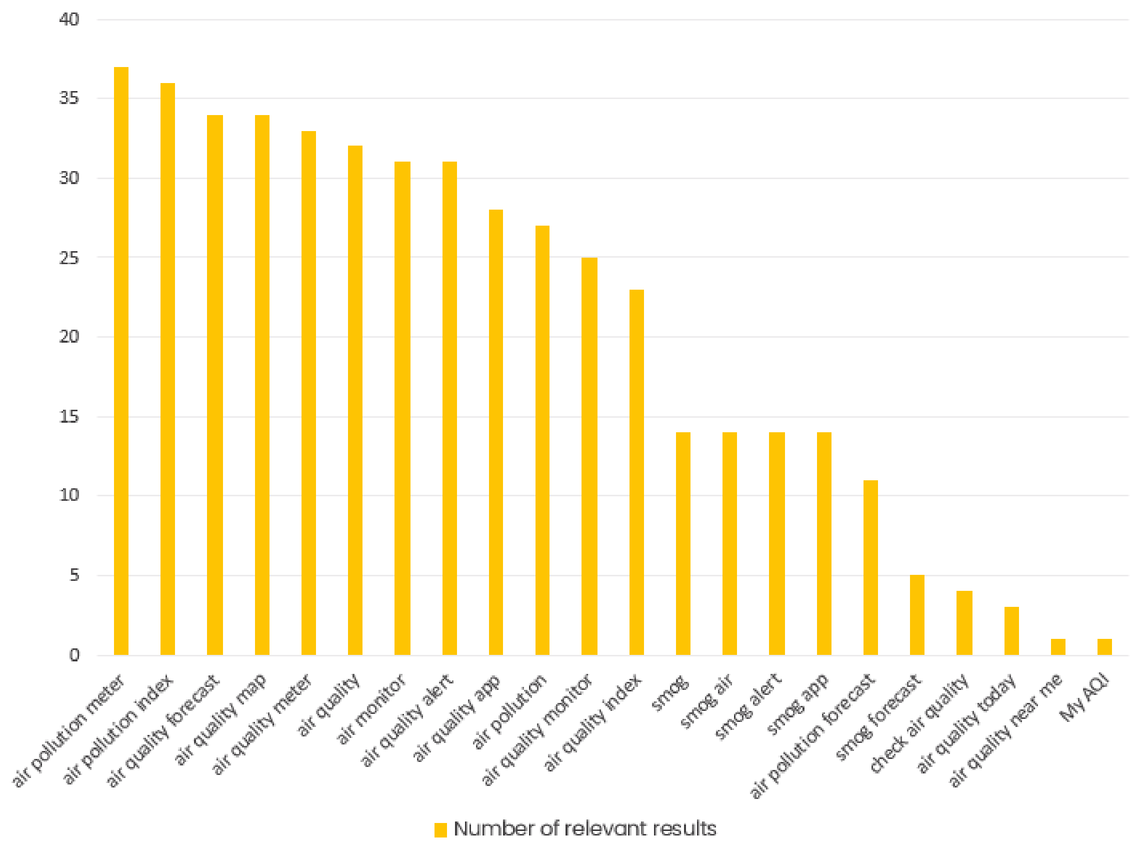
SerpAPI querying of the Apple App store resulted in a total of 1619 apps, which were refined to 146 after filtering for unique results.

Figure 33 Results returned via the Apple App store after the refining process



The search terms ‘air quality forecast’, ‘air quality alert, and ‘air quality map returned the highest number of relevant results from the Apple app store.

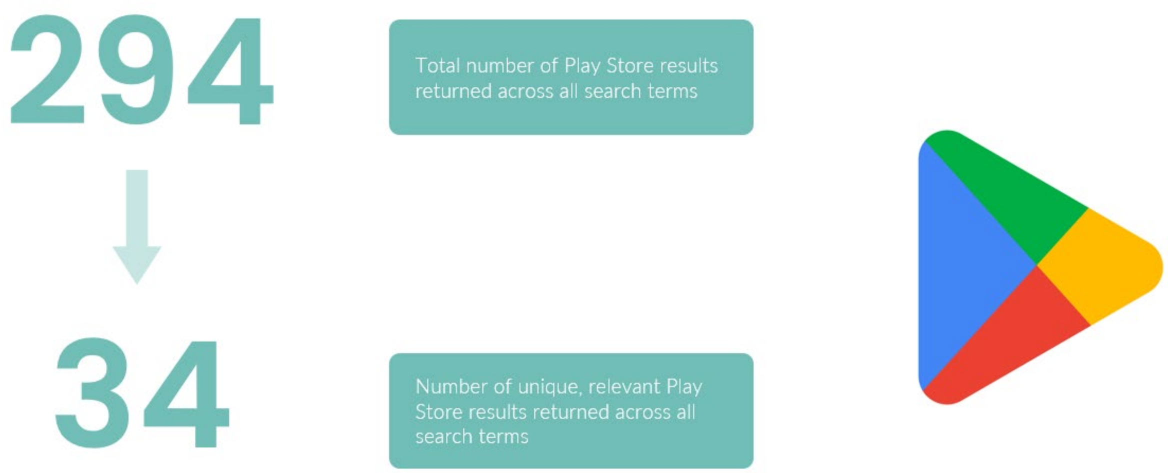
Figure 34 Relevant results returned from Apple iOS store for each key search term



9.5.2 Google Play Store

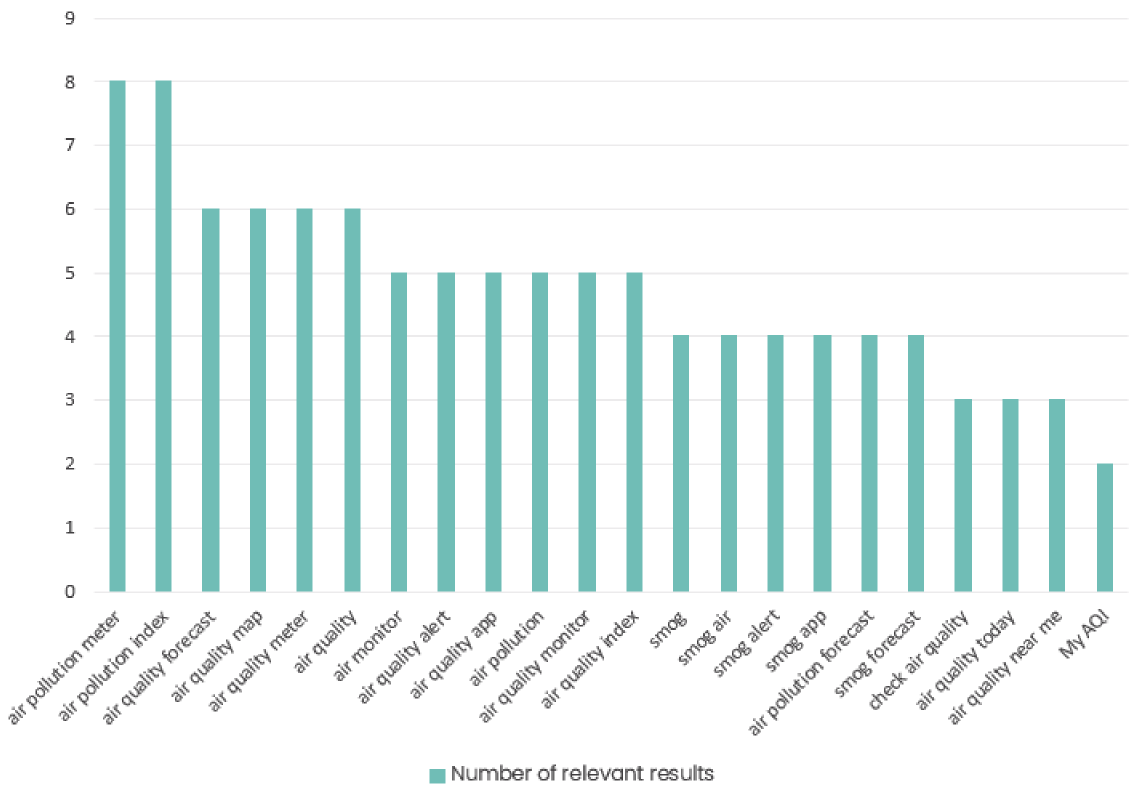
A total of 294 apps were returned from serpAPI querying of the Google Play store using the list of seearch terms. After filtering for duplicates and implementing manual checks, 34 relevant app-based AQ channels were retained.

Figure 35 Results returned via Google Play stores after refining process



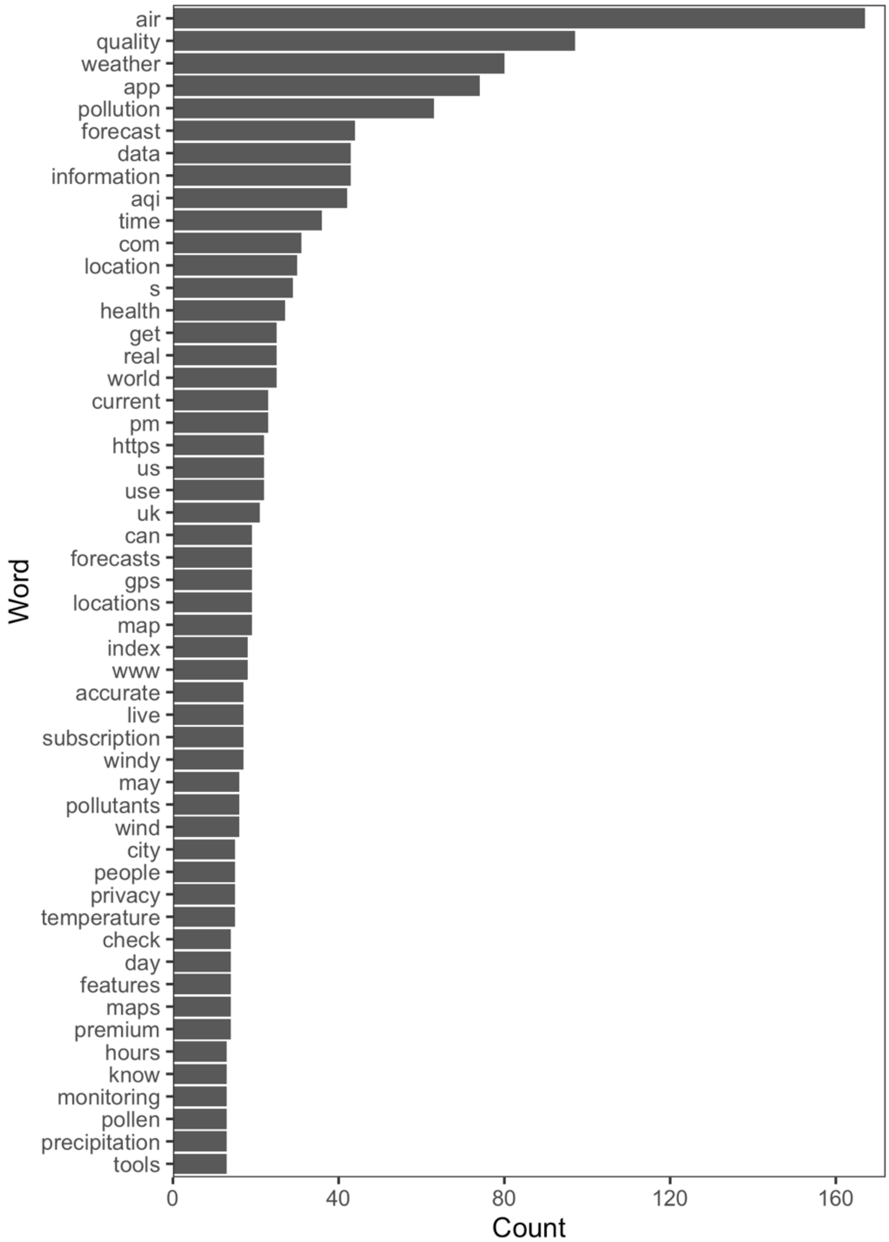
The search terms 'air quality', 'air quality index, and 'smog forecast' returned the highest number of relevant results from the Google play store.

Figure 36 Relevant results returned from Google Play store for each key search



9.6.1 Refinement process & common AQ channel descriptors

Figure 38. Common terms describing AQ channels



The most commonly featured terms across known AQ channels were used to help refine the dataset resulting from the systematic search (see Figure 38). The 52 search terms appearing most frequently in air quality related apps on the google playstore were loaded and for the html text corresponding to each unique url, and each unique domain, a 1 or 0 was recorded, indicating whether the html text did or did not contain each search term irrespective of case. A column was then added indicating the sum across all search terms for each unique url and unique domain.

To filter out duplicate domains (e.g. https://www.breathelondon.org and https://breathelondon.org/) domain which matched after the removal of the 'https://' prefix, and the 'www.' were identified and only one copy of the domain retained. Domains for which the suffix differed (e.g. https://www.air-quality.com and https://www.air-quality.org.uk) were retained as these reflect different websites.

10 References

[1] Schulte, K., 2022. 'Real-time'air quality channels: A technology review of emerging environmental alert systems. *Big Data & Society*, *9*(1), p.20539517221101346.

[2] di Bella E, Leporatti L, Maggino F (2018) Big data and social indicators: actual trends and new perspectives. *Social Indicators Research* 135(3): 869–878.

[3] Kenis, A. and Barratt, B., 2022. The role of the media in staging air pollution: The controversy on extreme air pollution along Oxford Street and other debates on poor air quality in London. *Environment and Planning C: Politics and Space*, *40*(3), pp.611-628.

[4] Marres, N., 2017. *Digital sociology: The reinvention of social research*. John Wiley & Sons.

[5] Bowker, G. C. (2005) Memory Practices in the Sciences, MIT Press, Cambridge, MA.

[6] Amoore, L. (2011). Data Derivatives. Theory, Culture & Society. https://doi.org/10.1177/0263276411417430

[7] Neff, G., Tanweer, A., Fiore-Gartland, B., & Osburn, L. (2017). Critique and Contribute: A Practice-Based Framework for Improving Critical Data Studies and Data Science. Big Data.

[8] Statista, 2023. Market share of leading search engines in the United Kingdom (UK) in April 2023. Retrieved electronically 20 June 2023. https://www.statista.com/statistics/280269/market-share-held-by-search-engines-in-the-united-kingdom/

[9] Gillespie, T., 2010. The politics of 'platforms'. *New media & society*, *12*(3), pp.347-364.

[10] SerpAPI, 2023. Retrieved electronically <https://serpapi.com/#header>

[11]Statista, 2023.Share of smartphone users in the United Kingdom (UK) 2012-2022, by age. Retrieved electronically https://www.statista.com/statistics/300402/smartphone-usage-in-the-uk-by-age/

[12] Statista, 2023. Most popular social media platforms in the United Kingdom (UK) as of the third quarter 2022, by usage reach. Retrieved electronically https://www.statista.com/statistics/284506/united-kingdom-social-network-penetration/

[13] Meltwater, 2023. UK Social Media Statistics. Retrieved electronically 20 June 2023 https://www.meltwater.com/en/blog/uk-social-media-statistics

[14]Mayor of London, 2023. Sadiq Khan issues ‘Very High’ air pollution alert in London. Retrieved electronically < https://www.london.gov.uk/press-releases/mayoral/mayor-issues-very-high-air-pollution-alert>

[15]Carrington, D. 2019. 'Location, location, lung disease': pollution ads target property market. The Guardian. Retrieved electronically <https://www.theguardian.com/environment/2019/may/01/location-location-lung-disease-pollution-ads-target-property-market>

[16] Little Black Book Online, 2020. Global Action Plan’s Real-Time ‘Breathable Billboard’ Urges Us Not to Let Air Pollution Go Back to ‘Normal'. Retrieved electronically <-https://www.lbbonline.com/news/global-action-plans-real-time-breathable-billboard-urges-us-not-to-let-air-pollution-go-back-to-normal>

[17] Cabinet Office, 2022. List of Cities. Retrieved electronically <https://www.gov.uk/government/publications/list-of-cities/list-of-cities-html>

[18] Google, 2023. How results are automatically generated. Retrieved electronically 30 May 2023 https://www.google.com/intl/en_uk/search/howsearchworks/how-search-works/ranking-results/

[19] Statista, 2023. Market share held by mobile operating systems in the United Kingdom (UK) from December 2011 to June 2022. Retrieved electronically 20 June 2023 https://www.statista.com/statistics/262179/market-share-held-by-mobile-operating-systems-in-the-united-kingdom/

[20] Sun, C., Kahn, M. E., & Zheng, S. (2017). Self-protection investment exacerbates air pollution exposure inequality in urban China. Ecological Economics, 131, 468–474. https://doi.org/10.1016/j.ecolecon.2016.06.030

[21] U.K. Department for Food, Environment and Rural Affairs (Defra), 2023. A Brief History of Air Quality Monitoring. Retrieved electronically https://uk-air.defra.gov.uk/networks/brief-history

[22] Zhang, Y., Bocquet, M., Mallet, V., Seigneur, C., & Baklanov, A. (2012). Real-time air quality forecasting, part I: History, techniques, and current status. In Atmospheric Environment. https://doi.org/10.1016/j.atmosenv.2012.06.031

[23] Kantar, 2018. This is how the UK uses social media
Retrieved electronically https://www.kantar.com/uki/inspiration/social-media/this-is-how-the-uk-uses-social-media

[24] Lauckner, C. and Whitten, P., 2016. The differential effects of social media sites for promoting cancer risk reduction. *Journal of cancer education*, *31*, pp.449-452.

[25] Ofcom, 2023. Media Nations 2023: Latest UK viewing and listening trends revealed. Retrieved electronically https://www.ofcom.org.uk/news-centre/2023/media-nations-2023-latest-uk-viewing-and-listening-trends-revealed#:~:text=As%20competition%20for%20the%20nation%27s,2021%20to%2079%25%20in%202022

[26] Kitchin, R., 2014. Big Data, new epistemologies and paradigm shifts. *Big data & society*, *1*(1), p.2053951714528481.

Imperial College London
Projects

Contact us:

Andrew Grieve, Environmental Research Group
Email: a.grieve@imperial.ac.uk

Dr Kayla Schulte, Environmental Research Group
Email: kayla.schulte@imperial.ac.uk

Imperial Projects
Email: iproj@imperial.ac.uk

Imperial Projects is a wholly owned company of Imperial College London

Covers designed using assets from Freepik.com