

# Mapping the news, media and journalism landscape

## Research report

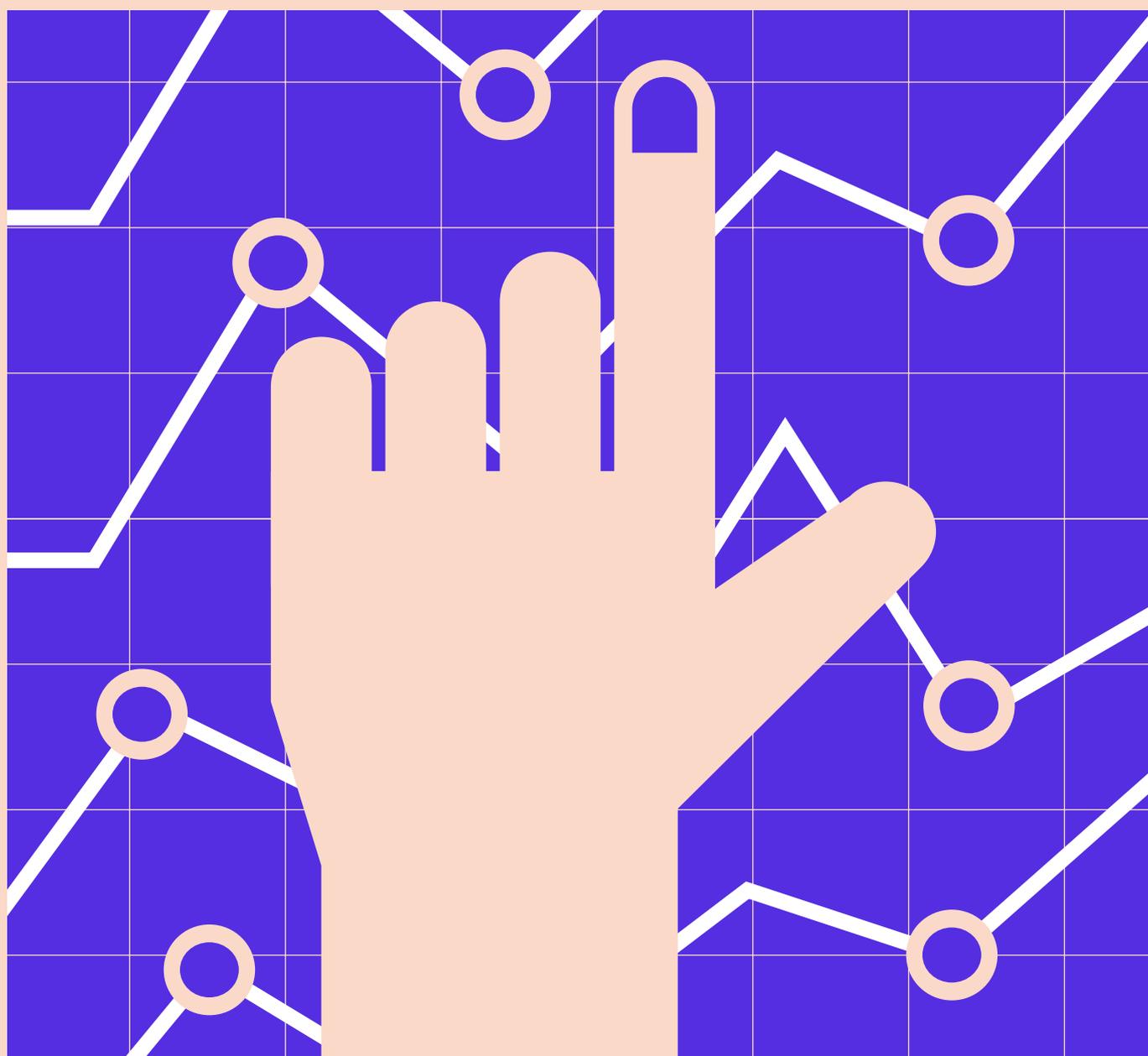
Adeola  
Otubusen

George  
Richardson

Chantale  
Tippett

Konstantinos  
Stathoulopoulos

Juan  
Mateos-Garcia



# Contents

---

<b>1.0 Summary</b>	<b>1</b>
<b>2.0 Structure of the report</b>	<b>3</b>
<b>3.0 Key findings</b>	<b>4</b>
<b>4.0 Background</b>	<b>8</b>
4.1 Context of this report	9
4.2 Industry background	10
<b>5.0 Research questions and methodological summary</b>	<b>13</b>
<b>6.0 Research findings</b>	<b>15</b>
6.1 The media and journalism landscape from multiple angles	15
6.2 Innovation in public interest news	44
<b>7.0 Conclusions and next steps</b>	<b>54</b>
<b>Acknowledgements</b>	<b>56</b>
<b>References</b>	<b>57</b>
<b>Appendix: Data, definitions, methods and limitations</b>	<b>58</b>
Datasets	58
Definitions	61
Methodologies	71
Limitations	72

# 1.0

## Summary

---

**Many industries transform or collapse in the face of mass disruption to business models and consumer habits. While often painful and challenging, such disruption can also lead to a vibrant new chapter in which more options and opportunities exist for all. Today, the news and journalism industry is undergoing a rapid transition, with incumbent players needing to adjust or perish in the face of new competitors such as Facebook and Google, new platforms such as the smartphone, as well as the new face of old issues such as the spread of misinformation and disinformation.**

While these changes offer windows of opportunity for new ideas and business models, the disruption of the news and journalism ecosystem also carries with it a unique risk. This is because news – in particular ‘public interest news’ such as investigative journalism and reporting on the daily activities of public institutions – is essential to the health of democracies. Although public interest news forms a small proportion of journalistic output overall, and is less in demand both from citizens and those who want to advertise to them than other kinds of journalism, public interest news is crucial for holding those in power to account and ensuring critical institutions in society (e.g. police, courts, health services) are functioning effectively.

In 2019, Dame Frances Cairncross was commissioned to examine the current and future market environment facing the press and high-quality journalism in the UK. In the final report, which ultimately focused on the need to safeguard and promote high-quality public interest news in the UK, Cairncross made several recommendations including the need to invest in more innovation in this fragile and increasingly threatened domain. It is within this context that Nesta launched the £2 million DCMS-supported Future News Pilot Fund, with the mission of providing grants and other support for innovation in public interest news in the UK. This study was undertaken to support a closely related end: to better understand the multiple dimensions of the media and journalism landscape in the UK, with a particular focus on public interest news.

The research presented here can be differentiated from previous research in this field in at least two ways: first, it draws together multiple data sources that tend to be analysed in isolation from one another (or not at all). Second, it leverages novel data sources such as Crunchbase in order to provide a deeper assessment of the public interest news innovation ecosystem.

Using official data from the UK’s Office for National Statistics, we find that news-related employment in the UK has remained relatively stable at around 85,000 jobs between 2011 and 2017, however the composition of employment has changed markedly. Employment in newspaper publishing and printing has experienced a rapid decline, while employment in web portals has increased (although in absolute numbers, newspaper publishing is still the largest employer). We also find that employment in journalism-related industries is very geographically concentrated relative to other sectors of the economy.

We then use data from Crunchbase to explore the startup landscape, finding a decline in the establishment of companies in news and journalism following a peak of activity in 2009. Although funding for news and journalism organisations has grown in absolute terms, it remains low, accounting for between 0.5 and 2 per cent of all funding in Crunchbase.

The data also suggest that news and journalism-related companies may be losing ground to growing sectors such as artificial intelligence (AI), blockchain and fintech. Relative to a set of comparator countries, the UK appears to have lower investment in news and journalism startups.

We then switch our attention to philanthropic funding activity, using two datasets with different geographic coverage (360Giving data for the UK and Candid data for the United States). Bearing in mind the limitations of these datasets and some challenges in comparability, we find that most philanthropic grants in news and journalism range between £6,000 and £62,000, and that the median grant size in the UK is 41 per cent lower than in the United States. The philanthropic funding landscape in the UK is also dominated by a small number of actors, with 73 per cent of UK funding coming from the top 3 funders, relative to 7 per cent of philanthropic funding in the United States coming from the top 3 funders.

Analysing the trends in academic outputs on misinformation and disinformation research using Microsoft Academic, we see an explosion of activity post-2016, likely as a result of the attention drawn to this subject during the 2016 US presidential elections. We find that while pre-2016 academic research on misinformation and disinformation is mostly related to medicine, post-2016 sees a radical shift toward research in politics, news and social media, internet privacy and fake news. Following the United States, the UK is the second most active country in misinformation and disinformation research, and 90 per cent of UK outputs come from institutions in London, Oxford and Cambridge.

In the final research section, we narrow in on the public interest news landscape in Crunchbase, building a picture of global activity, financial support and the adoption of technology. We focus on companies working in inclusive news, legal and civic news, high-quality news, news revenue and news engagement, finding that all sectors have seen a decline in startup activity as a share of news and journalism companies over the past two decades.<sup>1</sup> We find that while a quarter of companies work in more than one of the public interest sectors (e.g. legal and civic news and high-quality news), companies focused on news revenue tended not to overlap with other categories – a finding worth exploring in more depth. We also find a small but growing amount of crossover between certain domains of public interest news and emerging technologies such as AI and blockchain. Funding for public interest news ventures appears to be very low and reliant on a mix of funding sources. Finally, we find that the UK is under-specialised in inclusive news and news revenue, and that UK venture capital funding is lower than for similar companies elsewhere.

By pulling together various sources of data, this study aims to begin a process of breaking down silos and starting a collective movement towards the pursuit of a more holistic and cohesive view of the news, journalism and media landscape in the UK and globally. It is our hope that this pursuit will be continued beyond the Pilot Fund and further developed by others in the field. At the time of writing, the provision of public interest news was facing an existential threat as COVID-19 further exacerbates declines in advertising revenue and employment conditions for journalists. While we were unable to shift our research strategy to look in detail at how this crisis was impacting the landscape, we hope that these crucial questions can be explored at a later date.

<sup>1</sup> Definitions for these five sectors are provided in the Research questions and methodological summary section of this report.

# 2.0

## Structure of the report

---

The report is divided into several sections:

<b>Key findings:</b>	This section presents bullet points of key findings that emerged from our analysis. These provide readers with a snapshot of insights.
<b>Background and context:</b>	We explore how the media and journalism landscape has changed over time, and the challenges and opportunities this creates. We also describe the context in which this research was conducted.
<b>Research questions and methodological summary:</b>	This section provides a high-level overview of the research questions, definitions, data sources and methods used in the study. More detailed descriptions can be found in the Appendix.
<b>Research findings:</b>	<p>In this section, we present the findings from the research undertaken. It is divided into two subsections:</p> <p>The media and journalism landscape, which explores trends in employment, startups and philanthropic funding, as well as an analysis of misinformation-related research since 2000. The geographic focus of this section is primarily on the UK, however other geographies are included for comparison where data permit.</p> <p>Innovation in the public interest news landscape, which explores company formation globally and by subsector, as well as the use of emerging technologies such as AI and blockchain in public interest news companies and how startups are being funded.</p>
<b>Conclusions and next steps:</b>	We take stock of the findings and draw several conclusions. We also outline a series of next steps that can be taken by actors in the industry as well as researchers.
<b>Acknowledgements and references:</b>	In this section we acknowledge all of the people, organisations and work that directly and indirectly contributed to the research undertaken for this report.
<b>Appendix on data, definitions, methods and limitations:</b>	This appendix includes more detailed summaries of the data used in the study, as well as details on key definitions (e.g. which keywords are used to describe public interest news in our analyses) and methods used. This section also includes a description of the limitations in the analyses, including important gaps in the data that merit caution when interpreting results. <sup>2</sup>

<sup>2</sup> We also point out important limitations (e.g., notable gaps in data) throughout the report.

# 3.0

## Key findings

---

This section presents an overview of the key findings that have emerged from our research. For further details, please refer to the Research Findings section of the report.

### 01. The changing landscape of news and journalism employment in the UK

We start our exploration using official data<sup>3</sup> on employment trends in news and journalism in the UK, laying the groundwork for the sections that follow. Official data help to contextualise the novel analyses we perform in later sections, and are regularly used by industry and government decision-makers.<sup>4</sup> We find that:

- In 2017, there were 85,000 people working in journalism-related industries and this has remained relatively consistent from 2011. However, the composition of jobs has changed over this period, with employment in web portals growing while newspaper publishing and newspaper printing employment experience a rapid decline.
- Despite its relative decline in employment composition between 2010 and 2017, employment in newspaper publishing remained the largest employer (in terms of absolute numbers) in the journalism-related industries we studied.
- Journalism-related employment in the UK is very geographically concentrated. For instance, 57 per cent of employment in newspaper publishing is concentrated in 10 Local Authority Districts, which is 4 times more concentrated than employment in a composite benchmark of all other industries in the economy. All journalism-related employment has become more geographically concentrated since 2009 (with the exception of employment in web portals).
- Using data from the Local Authority Health Profiles published by Public Health England, we find that locations with lower levels of journalism employment tend to be less educated, older and more likely to have voted leave in the Euro referendum.

3 Complete descriptions of all data sources and methodologies used in this report can be found in the appendix of the report.

4 For instance, see Overview of recent dynamics in the UK press market (Mediatique 2018) and the Journalists at Work report commissioned by the National Council for the Training of Journalists (Spilsbury 2018).

## 02. The news and journalism startup squeeze

Following our analysis of employment in news and journalism, we shift our focus to a global perspective, investigating the number of private and public sector companies emerging in and around news and journalism.

- Using Crunchbase, a licensed database of public and private sector companies, we find that the number of companies being founded in the sectors of news and journalism has undergone both absolute and relative declines in the last decade. This follows a peak of activity around 2009, where 2.2 per cent of all companies were categorised as news organisations. This number has now fallen to 1.5 per cent.
- Funding for news and journalism has grown in absolute terms, but remained at a low and slightly fluctuating level since 2008, accounting for between just 0.5 and 2 per cent of all funding in Crunchbase.
- The levels of activity and funding are significantly lower than in younger but rapidly growing sectors such as AI, blockchain and fintech, suggesting that news and journalism is one of many traditional sectors losing market share of investment and funding to disruptive technologies.
- In the UK, funding for news and journalism companies has ranged from as little as \$5,000 to as high as \$80 million, but ranks in the lower half for average investment amounts when compared to other countries with a similar socio-economic profile and a significant presence in Crunchbase.
- The UK is slightly under-specialised in news and journalism, yet ranks in the top third for specialisation in the sector when compared to those countries.

## 03. Measuring philanthropic and nonprofit support for news and journalism

In this part of the analysis, we focus on funding for news and journalism that originates from philanthropic foundations, nonprofit organisations and civil society. The philanthropic side of funding on news and journalism projects with the caveat that, globally, data on philanthropic funding of journalism is still very uneven and of variable quality, and does not yet include data from major journalism funders such as Google News Initiative, Luminate or Open Society Foundations. We review the sector comparisons between the UK and United States and also take a deeper look into the sectoral and geographical differences within the UK.

- Using two datasets on philanthropic funding (360Giving for the UK and Candid for the United States), we identify 431 UK projects and 7,635 United States projects in news and journalism that have received at least \$5,000 (or the equivalent in GBP). We find that the majority of philanthropic spending in both countries falls between a lower quartile of £6,000 and an upper quartile of £62,000. However, the median grant size for projects in the UK is 40 per cent smaller than equivalent funds in the United States.
- From the 41 UK funders funding news and journalism in the data, the top 3 awarded 73.2 per cent of UK funds in news and journalism. In the United States, the top 3 funders awarded just 7 per cent of funding instances (valued over \$5,000).

- News and journalism make up a tiny fraction (less than 1 per cent) of nonprofit activity in the UK and represent a very small sector of activity in comparison to other social value sectors (e.g. sports, with a 13.9 per cent share of total UK projects in 2012).
- Nonprofit-funded news and journalism projects are somewhat distributed across the UK, but the largest area of concentration is London, with some smaller peaks of activity in other cities. We find that these metropolitan areas and some areas surrounding them exhibit specialisation in news and journalism.

#### 04. A change in the scale and scope of misinformation research

Since the 2016 US elections, there has been a growing discourse around ‘fake news’ and trust in the media. We analyse trends of academic research activity on misinformation, disinformation and fake news to understand how research interests have shifted between 2000 and 2019. We identify not only the volume of activity, but also the geographic and institutional distribution of publications using Microsoft Academic, an open research database.

- Research activity before the 2016 US Elections was mostly on misinformation in medicine. Post-2016, the focus shifted radically towards politics, internet privacy, and fake news.
- The UK produces the second largest volume of misinformation publications (after the United States). However, more than 90 per cent of the publications are from institutions based in England (specifically London, Cambridge and Oxford).

#### 05. Public interest news: Hints of slowdown and evolution

With an understanding that public interest news<sup>5</sup> captures a broad range of socially valuable news and journalism activities, we define five strands relating to public interest journalism or supporting sectors that are related to the DCMS/Nesta Pilot Fund. We use this to build a preliminary but nuanced global picture of the activity, financial support and technology proliferation within public interest news.

- We identify the volume of companies from Crunchbase working in inclusive news, legal and civic news, high-quality news, news revenue and news engagement (around 450, 450, 750, 1,800 and 2,300 respectively). All of these sectors have seen a gradual but steady decline in founding activity as a share of companies working in news and journalism over the last two decades.
- There are significant levels of overlap between public interest news themes, with around 25 per cent of companies identified as operating within two or more dimensions of five sectors. However, companies that are more focused on creating revenue for news organisations and developing revenue models are less likely to be involved in any of the other themes, indicating separate spheres of business activities.

<sup>5</sup> We acknowledge that the definition of what constitutes ‘public interest news’ varies amongst stakeholders. Below we specify how public interest news was defined in the Cairncross Review. In the Appendix of the report, we provide the keywords used for our analyses of public interest news.

- Public interest news sectors have received a very low (typically below 0.2 per cent) share of the funding and investment recorded in Crunchbase. There are some organisations that have received very large instances of funding (between \$100 million and \$1 billion), such as AOL Lifestream, Vice and Dataminr. Legal and civic news is the only sector that does not have any of these occurrences.
- Each public interest news sector relies on a mix of funding sources. In the UK, public interest news firms rely on angel investment, grants, product crowdfunding, equity crowdfunding and seed funding, and have received a lower proportion of venture capital than the global average.
- Investigating the specialisation of the UK in these dimensions of public interest news with respect to other countries reveals that the UK is under-specialised in inclusive news, slightly under-specialised in news revenue, moderately specialised in high-quality news and has proportionate levels of activity in legal and civic news and news revenue.
- There has been a growing crossover between technology and public interest news as measured by the number of companies operating across sectors. In particular we see a 10 to 20 per cent crossover between social media and the public interest news sectors, which is the same as or higher than for news and journalism companies in general. We also see a small but growing crossover between more recently emerging technologies, such as AI and blockchain, and certain sectors of public interest news.

# 4.0

## Background

---

**What purpose does news serve in society today? How equitable is the production of – and access to – news? In a rapidly evolving technological landscape, how can we verify the veracity and validity of the news we are consuming? When is government intervention in the news industry merited, and when does it threaten the independence of the sector?**

These are but a few of the important questions that have risen to the forefront of an increasingly urgent debate in the UK and elsewhere around the role of news in a rapidly changing world. Media business models (especially those reliant mainly on advertising) that have existed for decades are being disrupted by digitisation, with personal habits and preferences shifting alongside them. In parallel, we are witnessing a rise in political polarisation, populism, and the spread of misinformation and disinformation<sup>6</sup>. The world of news looks very different to how it did even a decade ago.

Although many industries transform or collapse in the face of mass disruption to business models and consumer habits, leading in many cases to a vibrant new era, such trends are particularly risky when it comes to news and journalism. This is because news – in particular ‘public interest news’ such as investigative journalism and reporting on the daily activities of public institutions – is essential to the health of democracies (Cairncross 2019). Although public interest news forms a small proportion of journalistic output overall, and is less in demand both from citizens and those who want to advertise to them than other kinds of journalism, public interest news is crucial for holding those in power to account and ensuring critical institutions in society (e.g. police, courts, health services) are functioning effectively.

Many already pressing questions about the role of news in the world have become increasingly urgent in the context of the COVID-19 crisis that was ongoing at the time of writing. Indeed, while questions of access to news and sustainability of public interest journalism were of primary importance at the outset of this work, the shock of the global pandemic has resulted in a more fundamental re-evaluation of how central news journalism is to our society, and has underlined how crucial it is that citizens be able to access trusted and verified information.

6 Disinformation is false information that is deliberately created or disseminated with the express purpose to cause harm, while Misinformation is information that is false, but not intended to cause harm (Wardle 2018).

## 4.1 Context of this report

In 2019, Dame Frances Cairncross<sup>7</sup> was commissioned to examine the current and future market environment facing the press and high-quality journalism in the UK. Cairncross worked with a small team of civil servants from the Department for Digital, Culture, Media & Sport (DCMS) to carry out the review, with guidance from an advisory panel.<sup>8</sup> The review considered a wide range of evidence, including an open call for evidence, input from stakeholders (e.g. publishers, advertisers, tech companies, academics) and commissioned research on the dynamics of the UK press market and online advertising in the UK.<sup>9</sup> In the process of collecting evidence for the review, the focus broadened from 'high-quality' to 'public interest' news.

The six-month Cairncross Review explored the questions of why we should care about the future of journalism, the changing market for news, the response of publishers to the shift online, the role of platforms in the markets for news and advertising, and the future of public interest news. The final report, published in February 2019, closes by asking what should be done, making nine recommendations ranging from commissioning a study on the inner workings of the online advertising market, improving news quality through regulatory means, exploring extending charitable status to journalism and implementing tax relief to support public interest journalism.

The report also highlights that addressing the broad swathe of challenges in the public interest news landscape in the UK will require a wide range of innovative responses – from new business models, to better data, to content that is more inclusive in both its production and consumption (e.g. across geographic, gender, race, age or other sociodemographic lines). To this end, Cairncross recommends that 'the government should launch a new fund focused on innovations aimed at improving the supply of public-interest news, to be run by an independent body' to be managed initially by Nesta before being transferred to a new Institute for Public Interest News (Cairncross 2019).

The initial government response to the Review was delivered by Conservative MP Jeremy Wright in a speech on 19 February 2019. Wright laid out immediate next steps including the intention to conduct a review on the regulation of online advertising and acknowledging the goal of 'driving innovation through a proposed new fund' (DCMS 2019). To this end, in January 2020, Nesta launched a £2 million pilot fund to provide grants and other support for innovation in public interest news in the UK. This study was carried out as part of the Pilot Fund in an effort to better understand the challenges and opportunities for public interest news innovation in the UK. The research team worked closely with the Fund's programme development and delivery teams to formulate the research questions and identify where the application of novel methods and analyses would add the most value to an already rich landscape of news and media research being undertaken in the UK by academia, industry and civil society.

7 Dame Frances Cairncross DBE, FRSE, FAcSS is a former economic journalist, author and academic administrator.

8 Including representatives from the fields of journalism, academia, advertising and technology. A list of advisory board members can be found on p.116 of the Cairncross Review (Cairncross 2019).

9 *Overview of recent dynamics in the UK press market* (Mediatique 2018); *Online advertising in the UK* (Adshead et al. 2019).

The research presented here can be differentiated from ongoing research in this field in at least two ways: first, it draws together multiple data sources that tend to be analysed in isolation from one another (or not at all), providing a rich picture of the UK landscape (benchmarked against other countries where appropriate or possible). Second, it leverages novel data sources such as Crunchbase in order to provide a deeper assessment of the public interest news innovation ecosystem. The use of novel data sources and methods allows us to answer questions that are challenging to do through the use of more traditional methods such as surveys. By pulling together various sources of data, the study aims to begin a process of breaking down silos and starting a collective movement toward the pursuit of a more holistic and cohesive view of the journalism and media landscape. It is our hope that this pursuit will be continued beyond the Pilot Fund and further developed by others in the field.

## 4.2 Industry background

Between 2007 and 2017, printed national newspaper circulation in the UK fell from 11.5 million to 5.8 million per day, and from 63.4 million to 31.4 million weekly for local newspapers (Mediatique 2018). By 2019, two-thirds of people were using a smartphone to access news weekly, and an increasing number of people had started accessing news in audio format (e.g. podcasts) or through online news aggregators (Newman 2019). This has drastic implications for how the news industry operates. Over the same decade that saw a halving of national and local print newspaper circulation, revenues from advertising and sales of printed newspapers also dropped by over 50 per cent, causing massive disruption to traditional business models (Mediatique 2018).<sup>10</sup> As news provision and consumption have moved to digital, online advertising has not been able to compensate for the decline in print revenue, leaving many publishers to develop coping strategies to try to fill the revenue gap lest they reduce or cease operations altogether. A number of strategies have emerged amongst news publishers trying to stay afloat: increasing the number of 'clicks' on their pages, increasing the value of advertising space on their page or attempting to get readers to pay for content.

The first strategy – seeking to maximise clicks that each only carry a low value – has proven insufficient for ensuring the financial success of most publishers, and critics have long argued that it may also incentivise 'clickbait'<sup>11</sup> (Cairncross 2019) and therefore erode trust. A submission to the Cairncross review highlighted that this approach also risks being particularly detrimental to the provision of local public interest news as local council reports will almost certainly draw fewer clicks than stories about 'WAGS'.<sup>12</sup> An alternative revenue generation strategy is to increase the value of advertising space by personalising ads to readers, however traditional publishers have struggled to compete against firms such as Facebook which hold vast quantities of data on what people like, do or buy (Cairncross 2019).

10 This situation has been drastically exacerbated by the COVID-19 crisis, with Enders Analysis predicting immediate and potentially existential threats in advertising and print circulation.

11 Overly sensationalist headlines for what are often lower quality stories.

12 An acronym used to refer to wives and girlfriends of high-profile sportspersons.

The third mechanism for trying to recover lost revenue is through direct payments from users,<sup>13</sup> which appears in some cases to work well for more established, global brands and less well for smaller brands (Cairncross 2019). More fundamentally, however, this shift requires a change in mindset, skills and approach to customer focus that has only recently started to be adopted in the field.<sup>14</sup> There are also striking differences between countries in people's willingness to pay for news. For instance, only 4 per cent of people in the UK are paid digital only subscribers,<sup>15</sup> relative to 8 per cent in the United States (US), 14 per cent in Sweden and 15 per cent in Norway. Additionally, if people are willing to pay anything at all for news, they are usually only willing to pay for one online subscription, leading to a 'winner takes all' effect (Newman 2019).

The shift to digital and disruption of traditional business models has also disproportionately threatened investigative journalism and democracy reporting, which are crucial to the effective functioning of society. For instance, reductions in democracy reporting is associated with a fall in community engagement by local residents, made visible through issues such as voter turnout and the management of public finances (Howells 2015). In conditions of financial uncertainty and low public demand for public interest news, incentives for producing this essential work have fallen. This has left specialist independent and nonprofit units like the Bureau of Investigative Journalism to take up the slack.

The shift to digital is not simply about the mode of content production and delivery – it has also brought a new form of competitor to traditional publishers. Indeed, the role of massive online tech platforms such as Google and Facebook have been central to the debate about the changing news ecosystem in the UK and elsewhere given their increasingly important role in the market for online ads and the distribution of news. A 2019 survey across nearly 40 countries found that Facebook products<sup>16</sup> reached 84 per cent of people, and 54 per cent of people indicated that they had used a Facebook product for news in the week prior to the survey (Newman 2019). Several challenges were highlighted in the Cairncross review with regard to the role these platforms play: the first is that the online advertising market where these players are dominant is confusing, particularly when it comes to 'programmatic display advertising'.<sup>17</sup> Second, as noted above, the sheer volume of data collected by the platforms makes it very difficult to compete with them. Third, the platforms can impose terms on publishers without consulting them, which threatens the viability of publishers' online businesses<sup>18</sup> (Cairncross 2019).

13 Referred to as 'reader revenue'. Modes of direct payment include subscriptions, memberships, donations and one-off payments.

14 See, for example, The Membership Puzzle Project: <https://membershippuzzle.org/>.

15 This measure aims to capture those who are prepared to use their own money to purchase a single title online news subscription. It removes those who get digital access with a print subscription or those who have a subscription paid by someone else. (Newman 2019).

16 Facebook also owns Instagram, Facebook Messenger, and Whatsapp.

17 The multi-step process of buying and selling digital ads.

18 For instance, publishers have expressed concern about the fact that the algorithms that determine how prominent a given news story is operate as a black box and can change without prior notice.

Amidst the challenges this new paradigm has created for publishers and journalists, we might have expected that this changing news landscape – in which a vast array of choices and personalisation are possible – would have at least made consumers happier or more informed. Evidence suggests that the situation is more complicated, however. Findings from the same multi-country 2019 survey cited above found that while people feel the news is good at keeping them up to date, they also feel that it is less good at helping them to actually understand a situation, and concern about misinformation/disinformation is high despite efforts by platforms/publishers to build public confidence (Newman 2019). Additionally, fewer than half (42 per cent) of people surveyed thought the media were doing a good job of holding rich and powerful people to account. More people are actively avoiding the news, citing the negative impact it has on their mood and the feeling of powerlessness it evokes – a trend that was most pronounced amongst people in the UK, who reported feeling sad, angry or bored with the topic of Brexit (Newman 2019).

In the UK, news produced may be more geared toward interests and needs of white, middle-class people living in urban areas, and those who are more educated (Newman 2019; Ofcom 2019; Spilsbury 2018). This may be due in part to who becomes a journalist in the UK. A 2016 study found that just over half of the UK's leading journalists were educated privately and 54 per cent went to Oxbridge (Kirby 2016).<sup>19</sup> The same study found some evidence of an improving situation, with just under 80 per cent of leading editors having attended either private or grammar schools relative to over 90 per cent in the mid-1980s. However, employment in the field appears to remain skewed in favour of elites.

Taken together, these trends suggest that although there has been a great degree of 'disruptive' innovation in the news industry, this may be having disproportionately negative impacts on the type of news needed to foster a healthy democracy. For instance, the role that Facebook and Google in particular play in the journalism sector is under constant scrutiny and controversy. While they provide enormous potential reach to audiences and the discoverability of content, as well as, latterly, direct grants to media organisations, they also take the lion's share of online advertising revenues, and have varied track records on matters related to freedom of speech, hate speech and censorship.

What solutions exist for addressing these challenges? The Cairncross Review highlighted several existing and proposed regulatory and policy options in the UK context, including tax relief aimed at encouraging payment for online news content and the provision of local and investigative journalism, as well as an easing of the rules for mergers of publishers experiencing financial distress (Cairncross 2019). Additionally, publishers have called for government intervention to ensure that the powerful position of platforms such as Facebook and Google do not cause undue harm to publishers.

A number of other solutions were explored in the Review, including customising stories to local areas, lowering the cost of local journalism, helping reporters to make better use of data, improving the digital infrastructure of public institutions (e.g. reliable wifi in courthouses) and ushering in new actors (e.g. Bureau Local in the UK). In response to people being worn out by the sheer volume and speed of news, we have also seen a rise in 'slow news' initiatives and constructive/solutions-based journalism (Newman 2019). Additionally, new models for explanatory journalism – which seek to explain complex issues in short videos – have become more popular.

19 Versus 93 per cent of the UK population on the whole attended state schools and 7 per cent who are educated privately.

# 5.0

## Research questions and methodological summary

---

**This report aims to answer a wide range of questions about the news and media landscape in the UK and, where appropriate, benchmark against other places in the world. Some of the key questions we address include:**

- How has the media and journalism jobs landscape changed over the years?
- Have jobs and businesses become more or less concentrated in places such as London?
- Have the absolute number of jobs changed, or just their location/focus (e.g. moving from print publishing to digital media)?
- What is the socio-economic profile of areas with low news activity?
- Who is funding innovation in public interest news?
- Where are new news and journalism companies being founded globally?
- How has the volume of academic research on misinformation, disinformation and fake news evolved over time?
- Which fields of study (e.g. medicine, political science) publish the most on these topics?
- What is the landscape of public interest news activity in sectors relevant to the Future News Pilot Fund?
- Which technologies are gaining traction in the sector?
- How is the UK faring in its support for new ventures in public interest news, particularly relative to other countries?

This section provides a brief methodological summary, including key definitions and points to bear in mind regarding data sources and methods. A more comprehensive description of all points covered here can be found in the Data, Definitions and Methods section of the report.

## Defining dimensions of public interest news

We understand that multiple definitions of what constitutes public interest journalism exist. For the data analysis in this report, we break the sector down into five more specific components. These are:

- **Inclusive news** – news relating to underserved or marginalised groups
- **Legal civic news** – reporting on legal and civic matters such as courts and government
- **High-quality news** – factual, unbiased and investigative journalism
- **News revenue** – activities that seek to generate income streams in the news sector
- **News engagement** – news that uses social media and other community engagement

These are defined by a set of data-driven keywords which can be found in the Definitions section of this report.

## Data summary

To provide a rounded picture of the public interest news sector both within the UK and a global context, we use data from several sources:

- **Business Register and Employment Survey (BRES)** – the official UK source of employee and employment estimates by geography and sector, collected by the Office for National Statistics.
- **Public Health England** – Local Authority Health Profiles published by Public Health England.
- **Crunchbase** – A global database of private and public sector companies, funding rounds and investors.
- **360Giving** – An open data platform collecting standardised grant data from over 120 UK nonprofits and foundations.
- **Candid** – A data aggregation and collection platform which collects philanthropic giving data globally, with a focus on data gathering from nonprofit digital tax filing in the US.
- **Microsoft Academic** – A widely used open database with more than 240 million academic documents (e.g. research papers and conference proceedings).

To analyse these datasets, we make use of open-source data analysis tools which form part of the Python scientific computing ecosystem. We also make use of well-known data science techniques, in particular for the identification of companies and projects relevant to public interest news with text data. The main method that we employ to achieve this is word vector keyword expansion. This takes a set of documents and builds a model that contains relationships between all of the words within the text. It is then possible to provide the model with a manually defined set of keywords in order to produce an extended list, which also contains many other related terms that have been found in the documents. This enables the creation of search queries using terms that may not have been initially apparent and that are definitely contained within the original documents. Although this method is data-driven, we acknowledge that it will not include all possible words relating to the sectors we are interested in, and that in some cases will lead to false positives. However, this is the case when searching with any large volume of data.

# 6.0

## Research findings

---

### 6.1 The media and journalism landscape from multiple angles

This section draws together findings from analyses conducted on multiple datasets: the Business Register Employment Survey (official data), private sector/startup activity (Crunchbase data), philanthropic funding data (360Giving and Candid data) and research activity (Microsoft academic). Each analysis sheds light on a different facet of the broader media and journalism landscape, and is intended to demonstrate what is possible when we start to break down the silos. This research study was part of a pilot fund, and was thus intended to lay the groundwork for broader and deeper analyses to follow. We pick up on the possible next steps in the Conclusions and Next Steps section of the report.

#### The evolving news and journalism employment landscape in the UK

In this section, we analyse data about employment in the UK's news and journalism sector, including trends over time, the composition of the sector, its evolution, geography and links with other important variables at the local level like economic, health, educational and political outcomes. This analysis sheds light on the disruption that the sector has experienced in recent years, and helps us to identify places most likely to experience under-supply of news (if we assume that locations that employ journalists are more likely to see their local news covered). It can also tell us something about the 'context' of locations lacking a journalism industry.

It is important to emphasise that these data capture the location of journalism employment, but tell us little about the extent and nature of the news being covered (i.e. if they are public interest news). Additionally, these data do not contain information on innovation, business models, barriers to growth and sustainability. Finally, the data are anonymised, making it impossible in most cases to dive deeper into the specific areas or questions of interest.

For the analysis of employment, we draw from the BRES (Business Register Employment Survey) data procured from the Office for National Statistics Nomis website.<sup>20</sup> We use the following standard industrial codes (SIC) to capture activity in news, media and journalism (which we refer to in the sections below as 'journalism'):

- 6012: Radio broadcasting
- 6020: TV programming and broadcasting
- 1811: Printing newspapers
- 5813: Publishing newspapers
- 6391: News agency activities
- 6312: Web portals<sup>21</sup>

<sup>20</sup> An analysis of establishment data from the Interdepartmental Business Register (not shown here) yielded qualitatively similar results to those on employment.

<sup>21</sup> Includes the operation of websites that use a search engine to generate and maintain extensive databases of Internet addresses and content in an easily searchable format, as well as the operation of other websites that act as portals to the Internet, such as media sites providing periodically updated content.

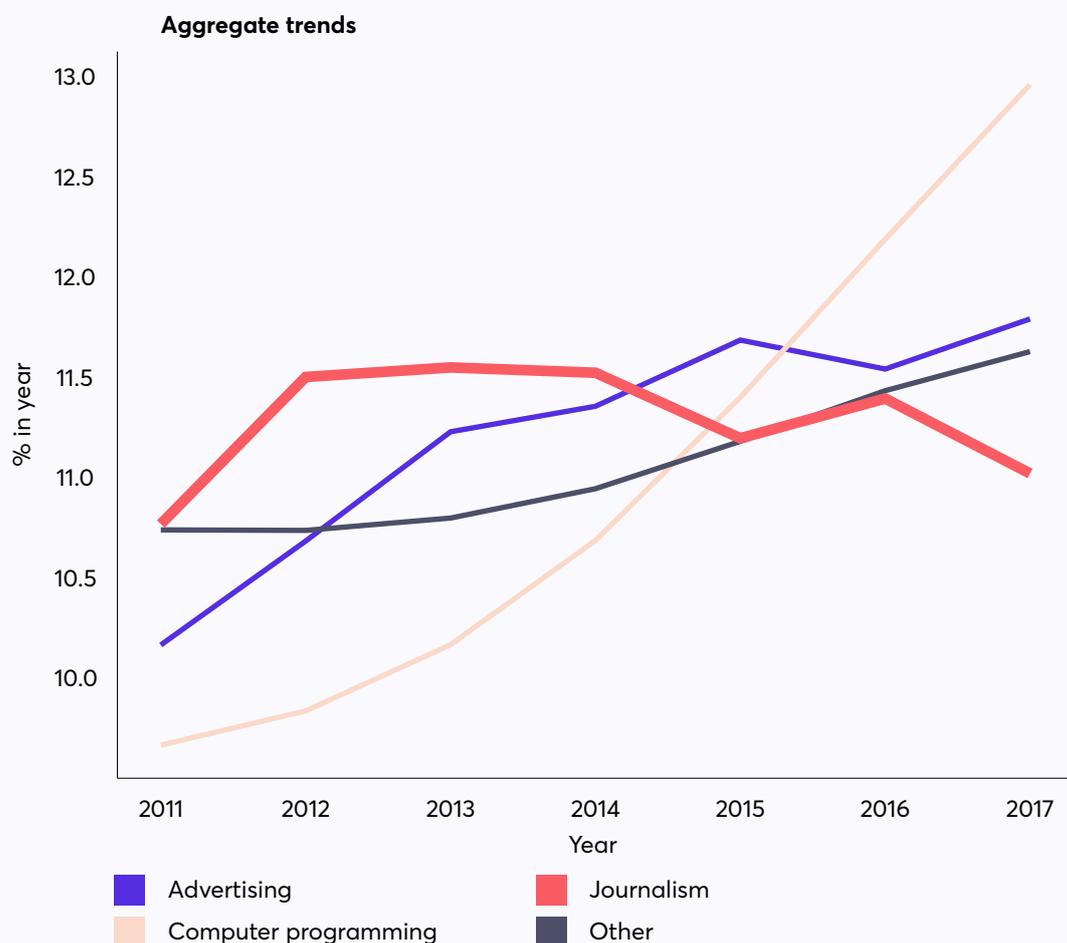
We benchmark activity in 'journalism' against activity in computer programming and advertising. The geographic unit of analysis is Local Authority Districts (LAD).

In addition to the BRES data, we use data from the Local Authority Health Profiles published by Public Health England to explore the association between journalistic activity and several demographic, health and economic outcomes. This component of the analysis (Figure 7) is restricted to England.

## Employment trends in journalism

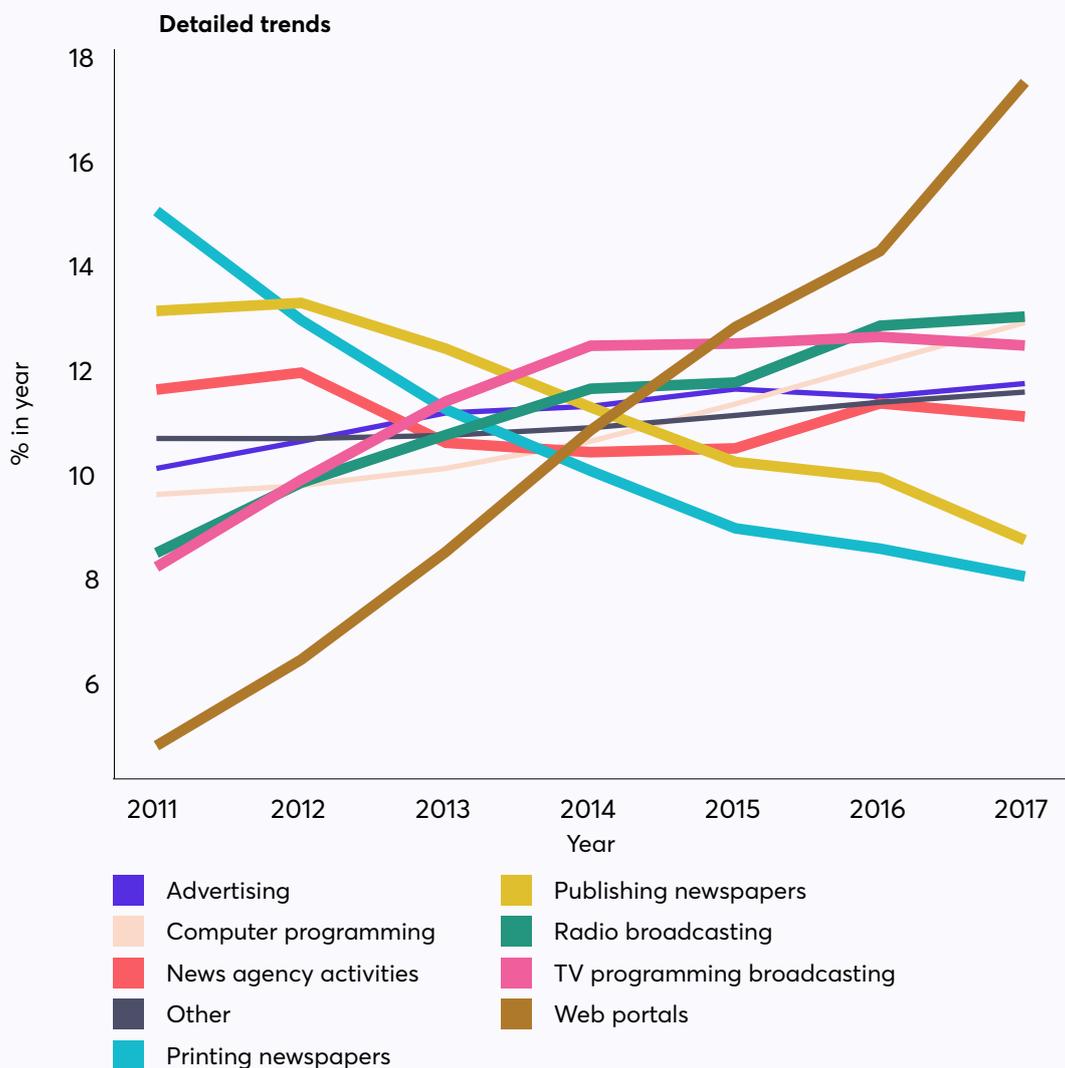
In Figure 1, we compare the percentage of all activity over a period accounted for by a single year in order to make it possible to compare the situation in sectors with very different volumes of employment. It shows that in 2017, there were 85,000 people working in journalism according to the BRES data. Despite some year-over-year fluctuations, journalism employment has remained relatively stagnant during the considered period (especially compared with computer programming, which has seen notable rise over the 2011–2017 period), with recent years accounting for a larger share of employment over the period. Employment in advertising was lower than in journalism until 2014, when employment in advertising surpassed it.

Figure 1: UK employment trends in journalism and benchmark sectors (2011-2017)



Aggregated trends presented in Figure 1 mask important changes in the composition of the sector. Figure 2 shows that web portal employment has been growing, as has employment in TV broadcasting (although with a recent plateau).<sup>22</sup> In line with trends highlighted in the Cairncross Review and elsewhere, employment in newspaper publishing and newspaper printing have been declining rapidly. Employment in newspaper publishing in 2017 was around half of what it was in 2009, for instance.

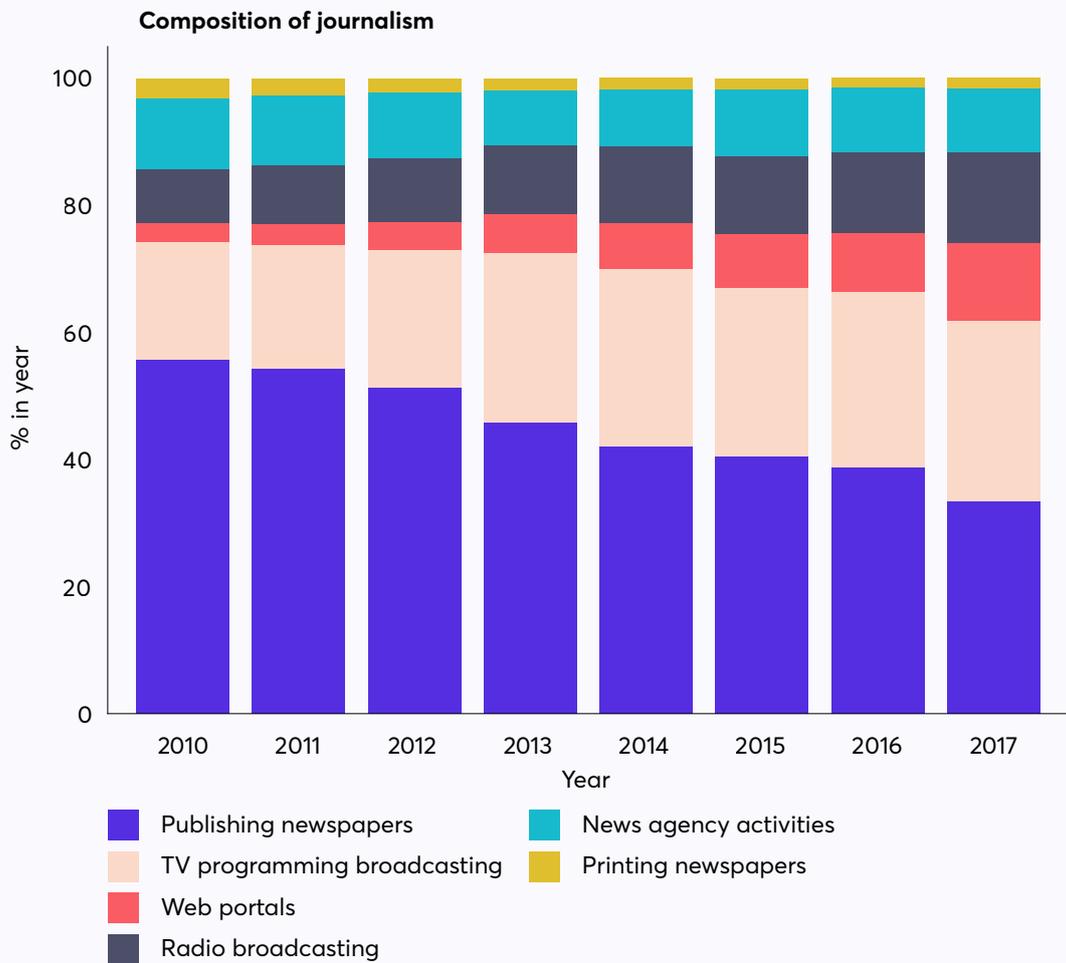
Figure 2: UK detailed employment trends in journalism and benchmark sectors (2011-2017)



As shown in Figure 3, although employment in newspaper publishing has represented a decreasing composition of the overall sector from 2010, in 2017 it remained the largest employer. Similarly, employment in web portals remains relatively low despite its rapid growth in recent years.

22 Although interpretation of this trend should also take into account the composition of journalism sector employment presented in Figure 3.

Figure 3: Composition of journalism sector employment (2010-2017)



## Geographic concentration of journalism employment in the UK

Journalism employment in the UK is very geographically concentrated compared to advertising, computer programming and other sectors<sup>23</sup> (Figure 4). For instance, 57 per cent of all employment in newspaper publishing is concentrated in 10 Local Authority Districts, which is around 4 times more concentrated than employment in other sectors. Web portals are even more concentrated, with the top 10 locations accounting for 68 per cent of employment.

Figure 4: Geographic concentration of employment in journalism and benchmark sectors

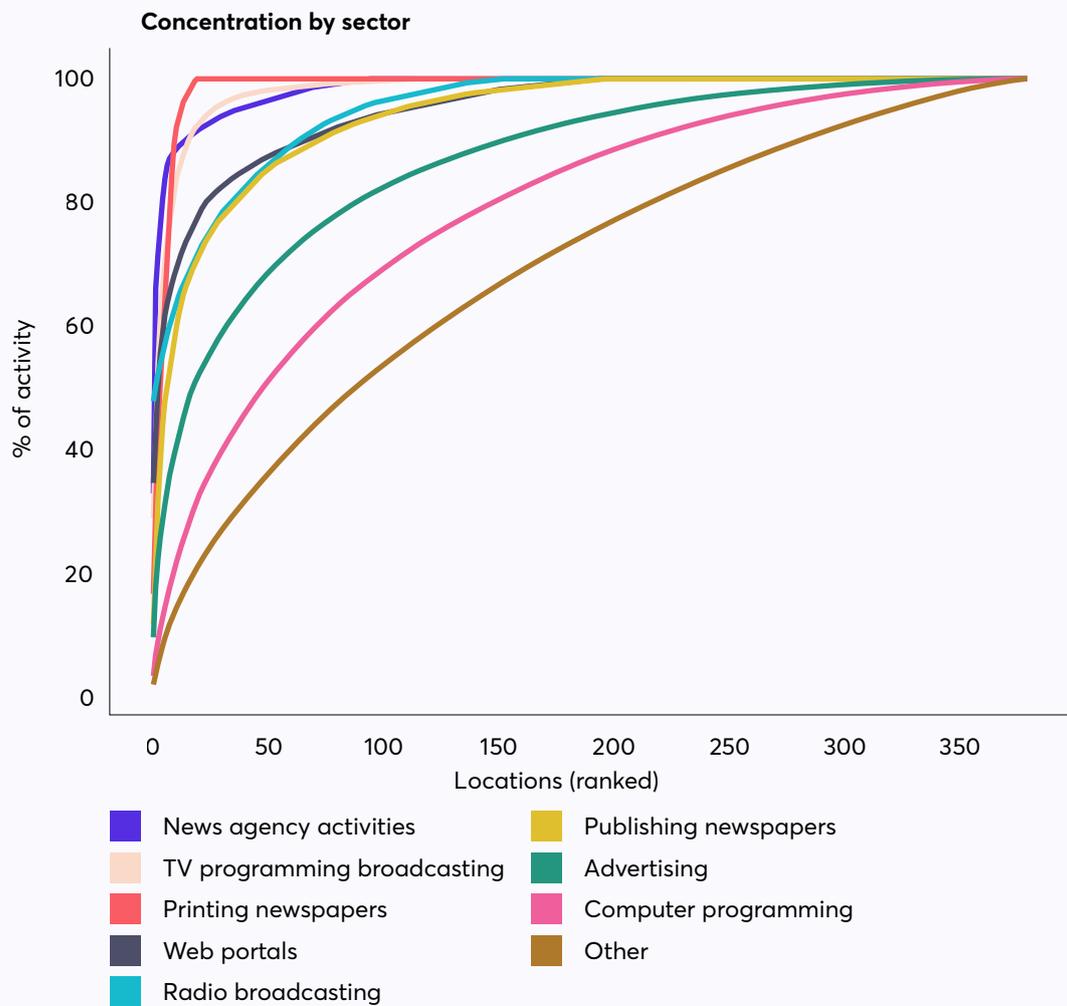
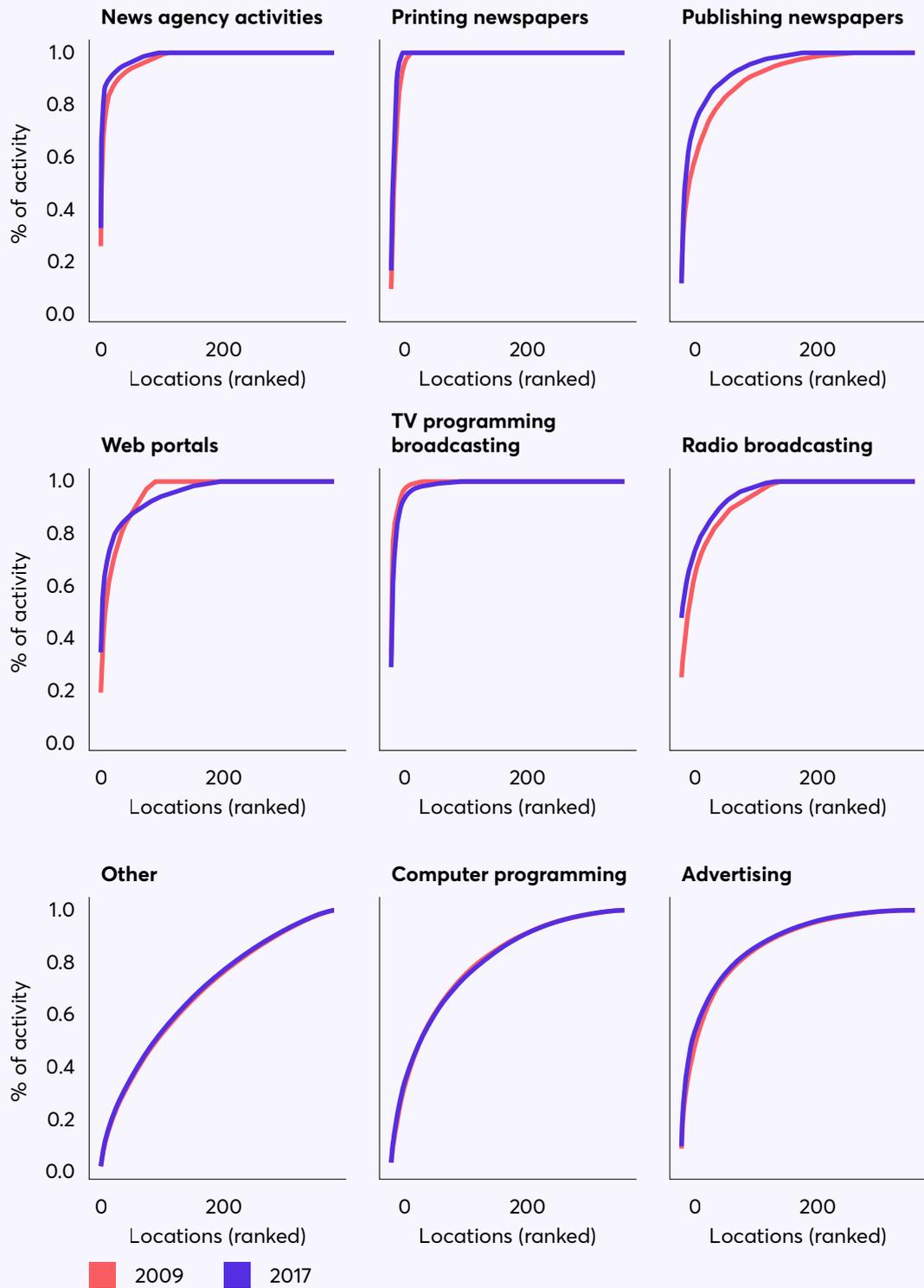


Figure 5 shows that all journalism-related activities became more concentrated between 2009 and 2017 with the exception of web portals, which became more evenly spread across the country (even though they are still highly concentrated in a small number of places).

Figure 5: Change in geographical concentration of journalism employment (2009 and 2017)



Relative to advertising and computer programming – which tend to cluster more strongly in London and the South – journalism employment hotspots are spread across the country (Figure 6a). This is consistent with the idea that there are concentrations of journalism employment serving the needs of regional audiences across the country. We see something similar when we compare newspaper publishing with web portals (Figure 6b). Note in particular the hotspots of web portal activity in the Midlands. Radio broadcasting is much less geographically concentrated than TV, possibly because there are less economies of scale and lower barriers to entry in this sector compared with TV broadcasting. This also suggests that radio broadcasting could play an important role in the diffusion of local news (Figure 6c).

Figure 6a: Geographical hotspots in journalism (right) employment relative to benchmark sectors of advertising (left) and computer programming (centre) (2017)

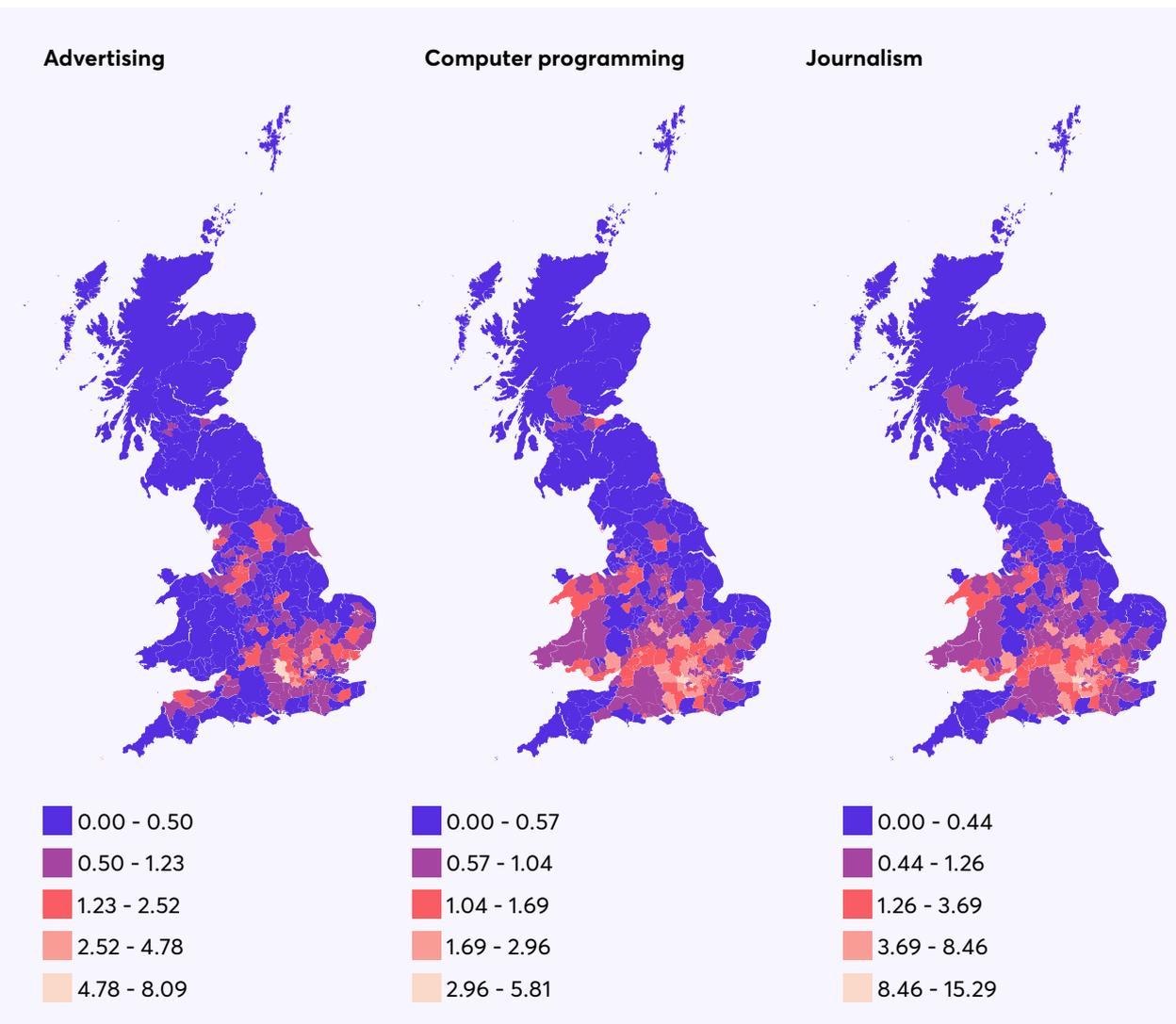


Figure 6b: Geographical hotspots in newspaper publishing (left) and web portal (right) employment (2017)

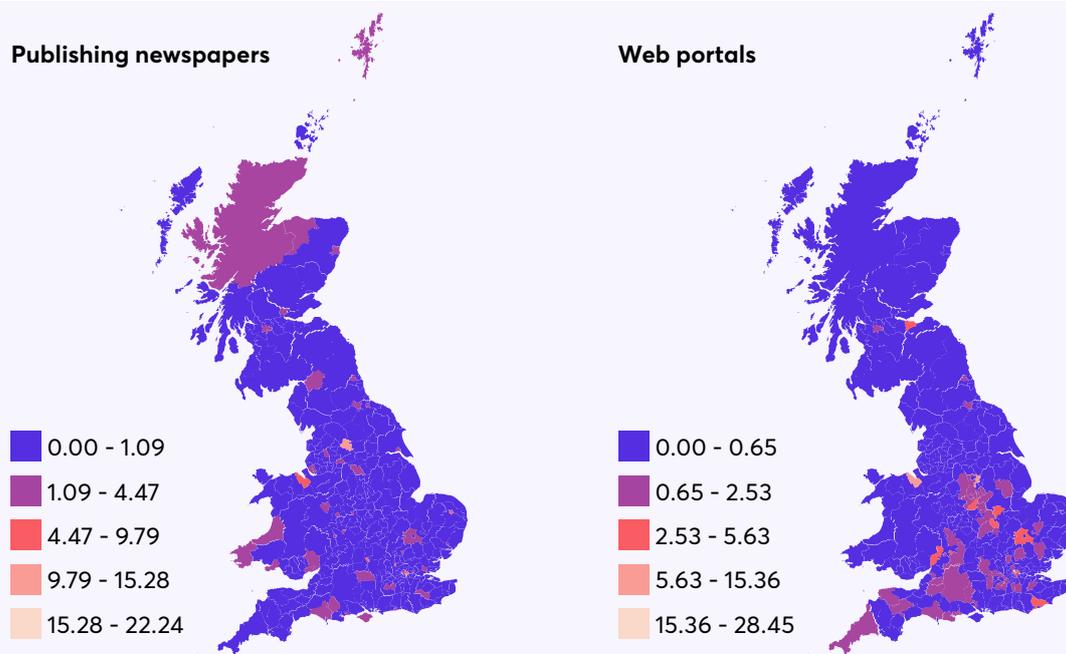


Figure 6c: Geographical hotspots in TV programming and broadcasting (left) and radio broadcasting (right) employment (2017)

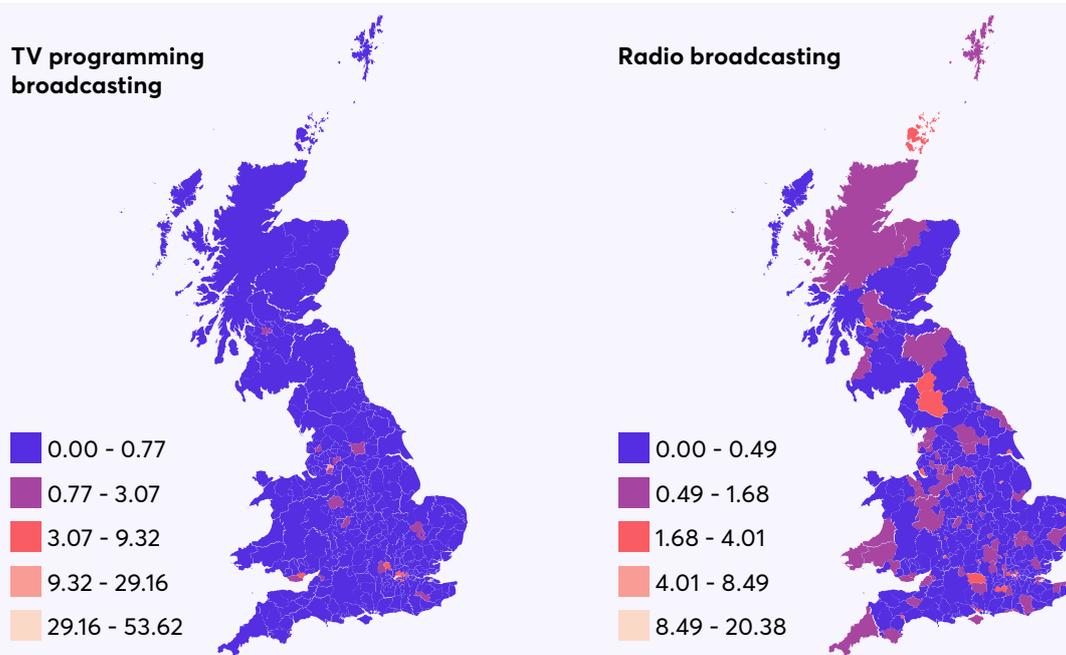
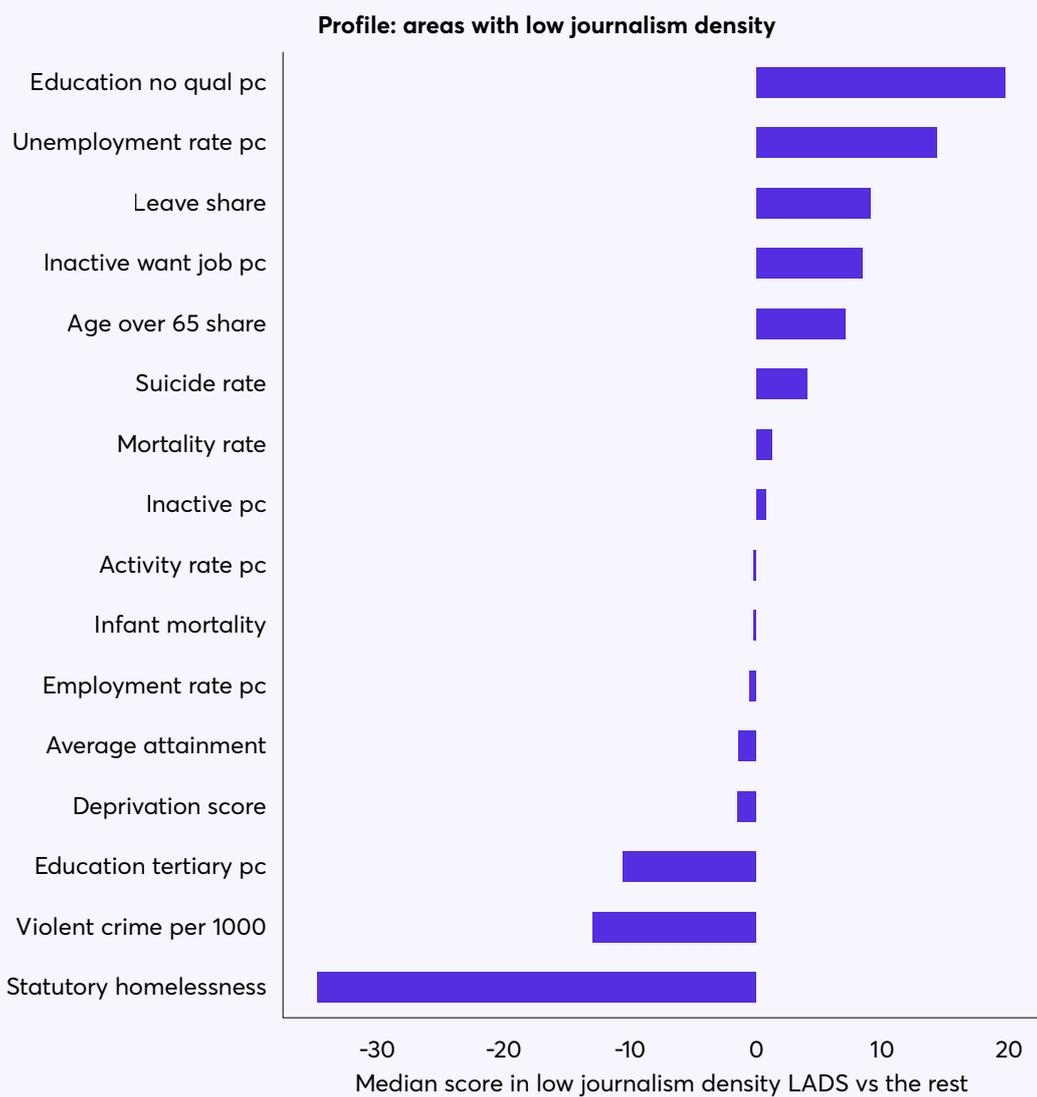


Figure 7 shows that locations with lower levels<sup>24</sup> of journalism activity tend to be less educated, older and more likely to have voted leave in the European Union membership referendum. They also have less crime and homelessness, which may reflect the fact that they are less urbanised.

Figure 7: Local profiles of locations in lowest quartile of journalism activity



24 Defined as Local Authority Districts in the bottom quartile of journalism employment

## Private sector activity and funding in news and journalism

Crunchbase is a database of over 700,000 private and public companies with global geographic coverage, with a particular focus on startups. The information it contains includes investments and funding, as well as industry category labels for each company. More information on Crunchbase can be found in the Appendix of this report. In the analysis presented here, we used the Crunchbase industry categories to identify organisations involved in news, journalism and social news. This resulted in 13,248 companies categorised as news, 579 categorised as journalism and 400 as social news.<sup>25</sup> We refer to these sectors together as news and journalism. We focus our analysis on companies founded between 2000–2018.<sup>26</sup>

We begin by situating the activity (number of companies founded) in news, journalism and social news companies within the larger Crunchbase dataset. The top plot in Figure 8 compares the sectors of interest to the growing sectors of AI, blockchain and fintech as a benchmark. Although news accounts for between just 1 and 2.2 per cent of companies founded each year, it is the 27th most active category in Crunchbase out of around 750. We can see that relative activity in news peaked around 2009, but this has since declined to 1.5 per cent, similar to levels around the year 2000.

As the graph shows the relative levels of activity when compared to the entire dataset, we should consider that in contrast, the peak of absolute activity was in 2013, when over 700 news companies were registered. In 2000 the number was only 141 and in recent years has declined again, with a mean of 487 companies being registered each year between 2015–2017, although this is partly due to a systematic lag in data collection by Crunchbase.

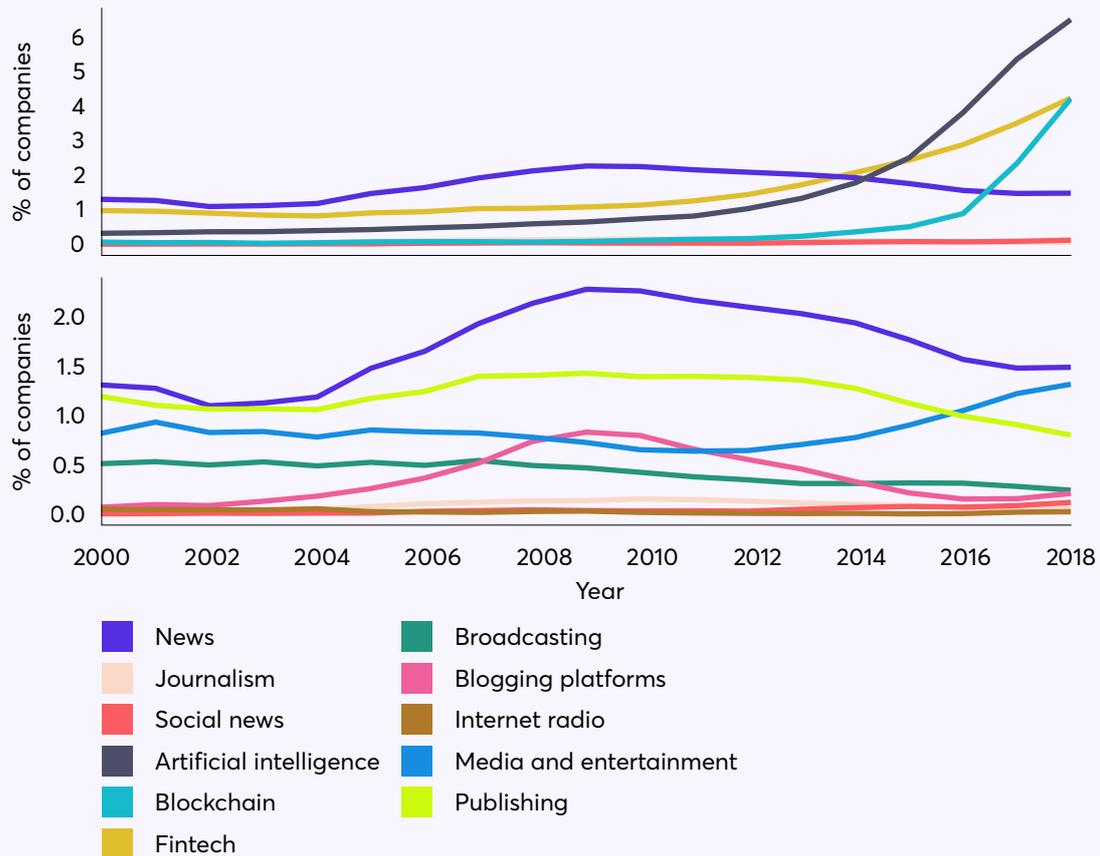
The benchmark sectors were chosen to highlight that trends differ across sectors. AI has seen both absolute and relative growth across the entire time period, accelerating rapidly after 2010. Such rapidly growing sectors may, in part, be responsible for the crowding out of more traditional industries such as news. Blockchain and fintech have experienced growth of 4.4 and 3.4 per cent respectively.

We have also compared news, journalism and social news to other sectors related to media, publishing and broadcasting in the bottom plot of Figure 8. Here it is easier to see that the level of activity in journalism has followed a trajectory consistent with that of news, but at a mean level of activity that is almost 20 times lower. Broadcasting has been experiencing a steady relative decline, while publishing and blogging platforms both experienced peaks in the late 2000s, since followed by a decline. Only media and entertainment has seen a significant level of relative growth in recent years perhaps accounted for by the rise in online streaming platforms that has emerged with the proliferation of high bandwidth consumer internet connections. Social news has seen some growth in recent years, presumably due to the rise of social media.

25 Company categories in Crunchbase are not mutually exclusive. For example, 229 companies are categorised as both news and journalism.

26 While the number of companies founded does not necessarily tell us how many companies are currently operating, only 1.1 per cent of companies are identified as having closed in Crunchbase, so we consider that the vast majority of companies here remain active in some form. Details on the data collection and curation methods used by Crunchbase to keep their database up-to-date can be found in the Appendix of this report.

Figure 8: News, journalism and social news company founding activity in Crunchbase (three year rolling mean).



The levels of activity referred to above are global figures, but there are important geographic differences that can be found in the Crunchbase data for news and journalism. To study this, we compared the UK to a set of countries that each have at least 30 news and journalism organisations on Crunchbase and that score above the median on the 2019 World Press Freedom Index.<sup>27</sup> This subset of countries was chosen to ensure that the countries being compared are more likely to contain a significant number of organisations who play a role in representing the public and holding institutions to account.

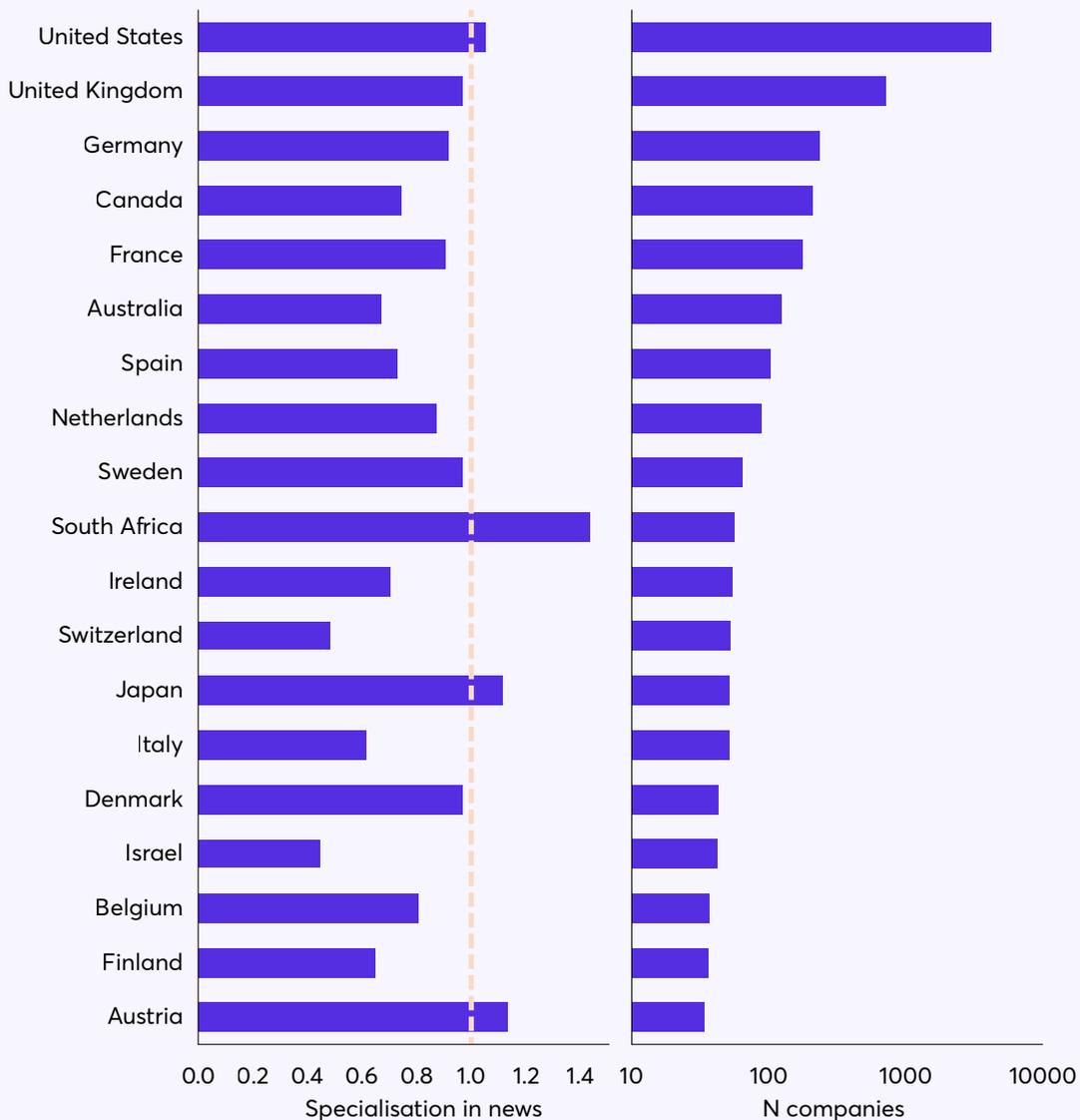
As Figure 9 shows, almost 4,200 of all the news and journalism companies identified were founded in the US, while just over 700 were founded in the UK. This accounts for 30.5 per cent and 5.2 per cent of the global count, respectively, while the other 17 countries account for 10.7 per cent. The stark difference between the US and the rest of the world suggests some potential systematic biases in the data collection methods used by Crunchbase. For this reason, we also show the specialisation (or location quotient) of the same countries in public interest news. The specialisation measure is a comparison between a country’s share of news organisations to its share of all organisations, thus identifying whether a country has a greater or lower than proportionate share of global news activity within its borders.

27 The 2019 World Press Freedom Index compiled by Reporter Without Border (RSF) obtained from [https://rsf.org/en/ranking\\_table](https://rsf.org/en/ranking_table)

A value of one indicates that the country has the number of companies we would expect based on its global share of all companies. A value greater than one indicates a degree of specialisation and a value between 0 and 1 indicates under-specialisation.

We can see that the UK is slightly under specialised in news, ranking sixth out of the 19 countries shown, with a specialisation value similar to that of Germany, Canada, France, Sweden and Denmark. Only the US, South Africa, Japan and Austria show some specialisation. This shows that although the UK is home to a large number of the public interest news companies listed on Crunchbase, there is some evidence that it is underinvesting in the sector relative to other industries, a finding that is in line with investigations into the UK’s funding landscape presented later in this report.

Figure 9: Specialisation in news and journalism (left) and number of news and journalism companies founded in the UK and comparison countries (right). The number of companies is represented on a logarithmic axis. Just over 30 were founded in Austria and almost 4,200 in the US.

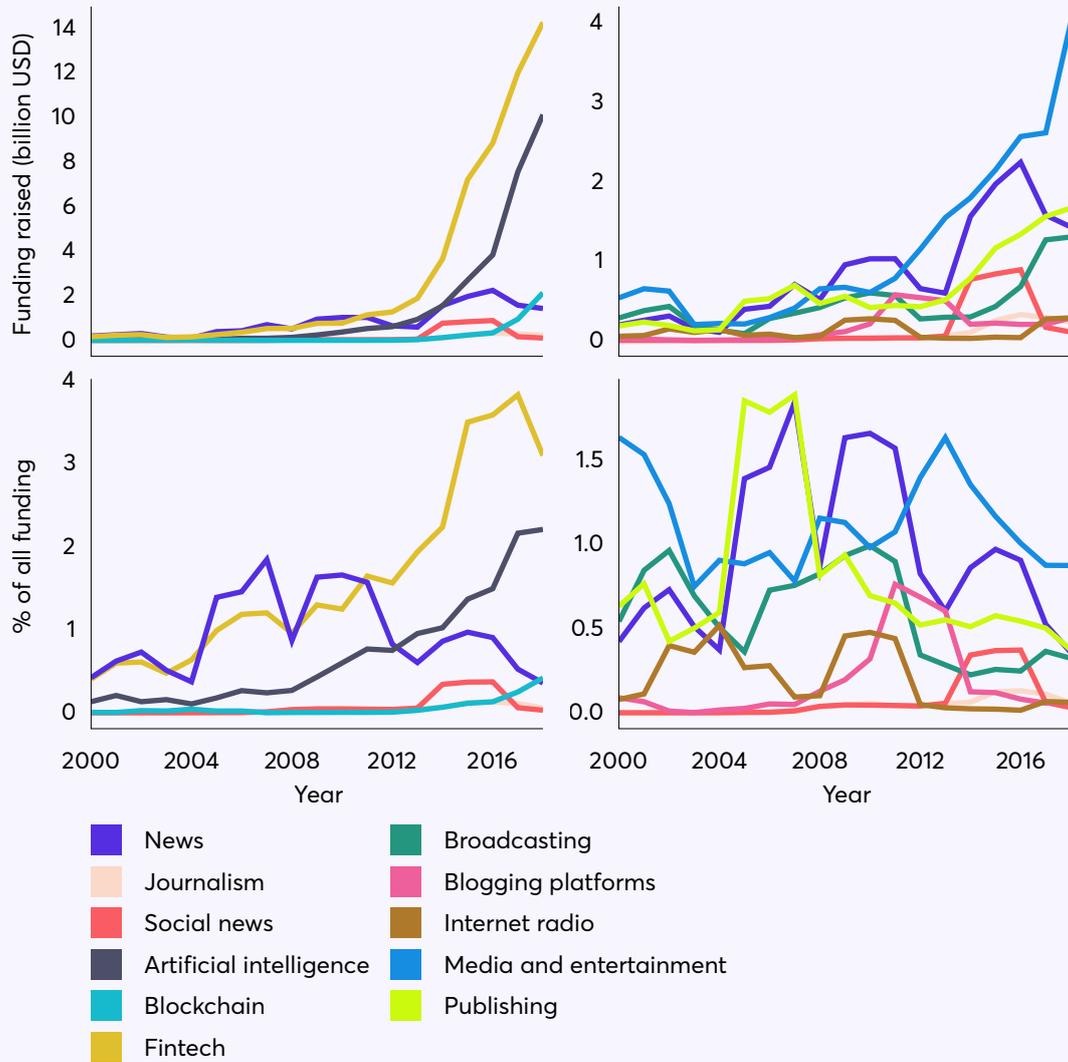


In addition to measures of activity, we consider the levels of funding received by those same news and journalism companies. Crunchbase contains information on the amount of money raised by companies from investment rounds, grants, seed funding, and more. The findings in this section are limited by the fact that funding data is only available for 12.5 per cent of news and journalism companies (lower than the Crunchbase average of 19.0 per cent). This is due to a variety of reasons including limitations in Crunchbase's data collection methods such as reliance on crowdsourcing, and the fact that some sources of funding are not publicly disclosed. Despite this, there is a sufficient volume of data to cautiously determine some trends and differences between sectors and countries in the funding landscape for news and journalism.

Here, the funding for news, journalism and social news companies has been compared to the same benchmark sectors as in the previous chart. Figure 10 provides an overview of the trends in funding between 2000–2018, with the top plots comparing the absolute funding between sectors. While almost all sectors have seen growth in the levels of investment, the increased funding and the rate of growth for news companies has been moderate. In comparison, the funding for AI in recent years has undergone a rapid increase resembling exponential growth, overtaking news in 2017 and 2018. When this is normalised by the total amount of funding raised each year across all companies in Crunchbase, we can see that news companies have accounted for around 1.5 per cent of all funding since 2005. However, between 2008 and 2018, the proportion of funding for the news sector fell by 0.27 percentage points, which must be considered against the increase in funding in growing sectors such as AI, blockchain, fintech (2.9 per cent, 1.1 per cent and 6.0 per cent respectively in the same period). The total funding for journalism and social news remains very low, with respective maximum shares of 0.2 and 0.9 per cent within the time period shown.

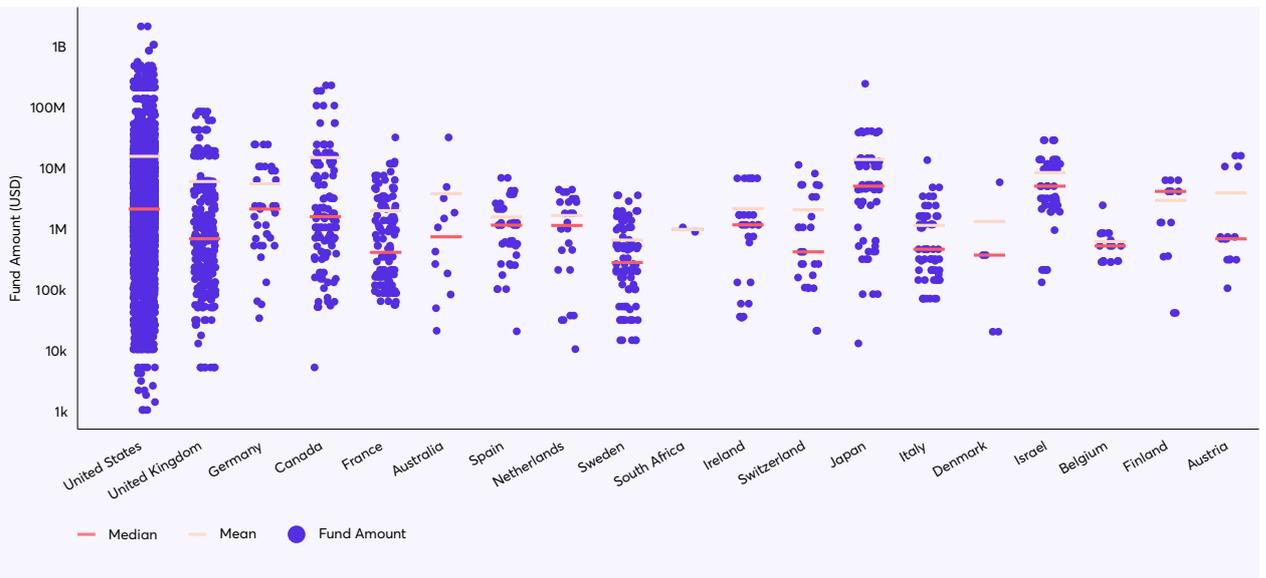
A more nuanced picture emerges when comparing news, journalism and social news to other similar sectors. While almost all media sectors shown have experienced a rise in absolute funding, in relative terms, they all account for less than 4 per cent of all funding in Crunchbase within a given year. What is clear is that news and publishing, which both experienced relative peaks around 2005, have since declined, while broadcasting and media and entertainment appear to be enjoying relatively high levels of funding in more recent years. Indeed, media and entertainment appears to have experienced steady linear growth over the last two decades, again highlighting the rise of online streaming subscription services.

Figure 10: Funding for news, social news and journalism in Crunchbase (three year rolling mean).



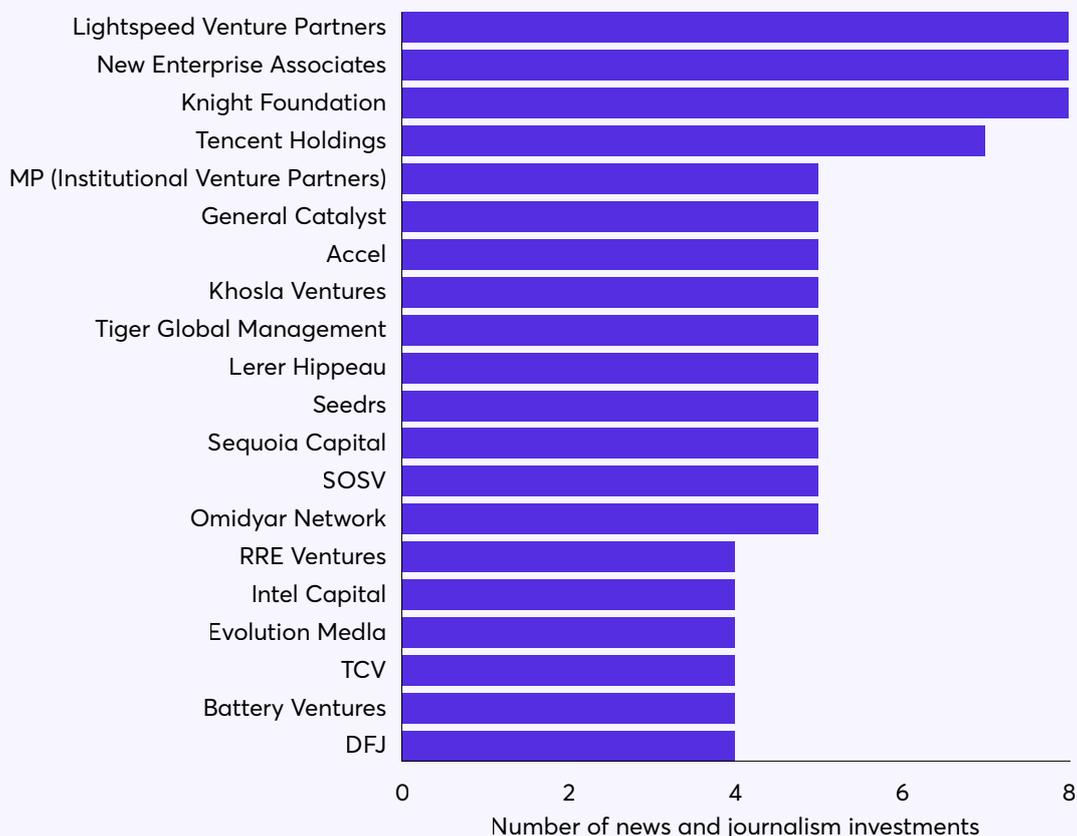
There is also an underlying geographical split within the funding trends for companies classed as news and journalism organisations. Looking again at the UK and the 18 comparison countries in Figure 11, we can see that the majority of funding falls between a wide range of \$10,000 and \$10 million (USD), with median values in the high hundreds of thousands or low millions. Only Canada, Japan and the United States have companies which have received funds of more than \$100 million, and only the United States has seen individual funds made to news and journalism in excess of \$1 billion. Overall, the UK ranks 12th out of 19 countries in terms of the median fund amount, but fifth when considering the mean. This indicates that although there have been a number of high value funds to companies in the UK, they generally receive less funding than their counterparts in comparison countries.

Figure 11: Distribution of funding in news for the top comparator countries. Note the logarithmic scale for the fund amount.



According to data available in Crunchbase, the majority of investors who have funded news and journalism companies are venture capital firms. Figure 12 shows organisations that have invested in four or more companies. Notable exceptions are the Knight Foundation and the equity crowdfunding platform, Seedrs.

Figure 12: Top global funders and investors of news and journalism



### Philanthropic funding in news and journalism

The development of news is of course not solely the domain of the private sector. Foundations, civil society, non-profits, and grassroots organisations also play a role in carrying out and funding news and journalism activity (Padania 2018). They may provide outputs in areas of significant public value and serve under-represented audiences that for-profit entities may overlook due to the lack of financial viability. Data about non-profit funding therefore captures dimensions of support for the news ecosystem and social innovations which are poorly represented in the Crunchbase dataset, which captures primarily for-profit and technology-oriented activity.

In this section, we look at the levels of philanthropic funding and activity in news and journalism to improve our coverage of relevant activities. We draw comparisons between the situations in the UK and the United States and analyse the concentration of funders and recipients, the geographical spread of activity and the distributions of activity between news and other non-profit sectors within the UK. The data on UK activity was obtained from 360Giving, a data aggregator of philanthropic projects in the UK that covers the period ranging from 1991– 2020. For the United States, data was licenced from Candid, another data aggregator of philanthropic projects, that specialises in collecting philanthropic giving data from digital tax filings in the United States. The Candid data spans the years 2010–2020.<sup>28</sup>

28 These two datasets have important differences in their methods of collection and coverage of relevant activities. This is explained in detail in the Appendix of the report.

To identify news and journalism organisations from 360Giving, a word vector keyword expansion was used, with initial news and journalism related seed terms selected from project descriptions in the dataset.<sup>29</sup> To the extent possible, we try to minimise the proportion of non-news and journalism related projects in the filtered data, however there is inevitably some variation in how closely related all projects are to the sectors of interest. Additionally, we removed unrelated projects that could skew the analysis.

For example, the charity Mind (formerly The National Association for Mental Health) based in the UK was awarded £5 million and included the sentence “good news piece” in the project description, but is not a news and journalism project. After cleaning the data, we were left with 431 news projects funded by 41 funders (out of a total of 126 funders).

The initial dataset obtained from Candid was a subset of their full database, including only projects related to Information and Communications. We filtered the data based on Candid’s philanthropy taxonomy, selecting projects which are labelled with the categories that are most related to news and journalism. The categories that we have used to extract the most relevant projects for news and journalism are:

- News and public information
- Journalism
- Advocacy journalism
- Citizen journalism
- Constituency journalism
- Investigative journalism
- Photojournalism

The resulting Candid dataset includes 8,472 projects funded by 1,435 organisations.

### **Funding bodies and recipients**

In the first part of the analysis we narrow the field of news and journalism projects to be compared between the UK and United States to those awarded at least \$5,000 (USD or the equivalent in GBP).<sup>30</sup> When comparing data from Candid and 360Giving, we use data from 2010–2019 as this is the overlapping period between the two datasets. This leaves us with 233 UK projects and 7,635 United States projects in news and journalism.

The Big Lottery Fund is the UK’s leading funder, having awarded funds to 127 recipients (Figure 13). The Wellcome Trust is the second largest funder with 31 projects. All other funders in the UK dataset have each funded fewer than 15 projects. In Figure 13, the United States funding environment appears to be much less concentrated in terms of the number of awards made by each funder. The top three UK funders make up approximately 73.6 per cent of the funds granted during this period, whereas the top three United States funders awarded just 7 per cent of United States funds for news and journalism projects.<sup>31</sup>

We have also compared the available institutional profiles of beneficiaries in the United States and UK and found that the top 10 recipients (in terms of the number of grants received) in both countries are mainly non-profit organisations in reporting and journalism

29 The word vector keyword expansion method is explained in the Methodologies section of the Appendix, and the final set of keywords can be found in the Definitions section of the Appendix

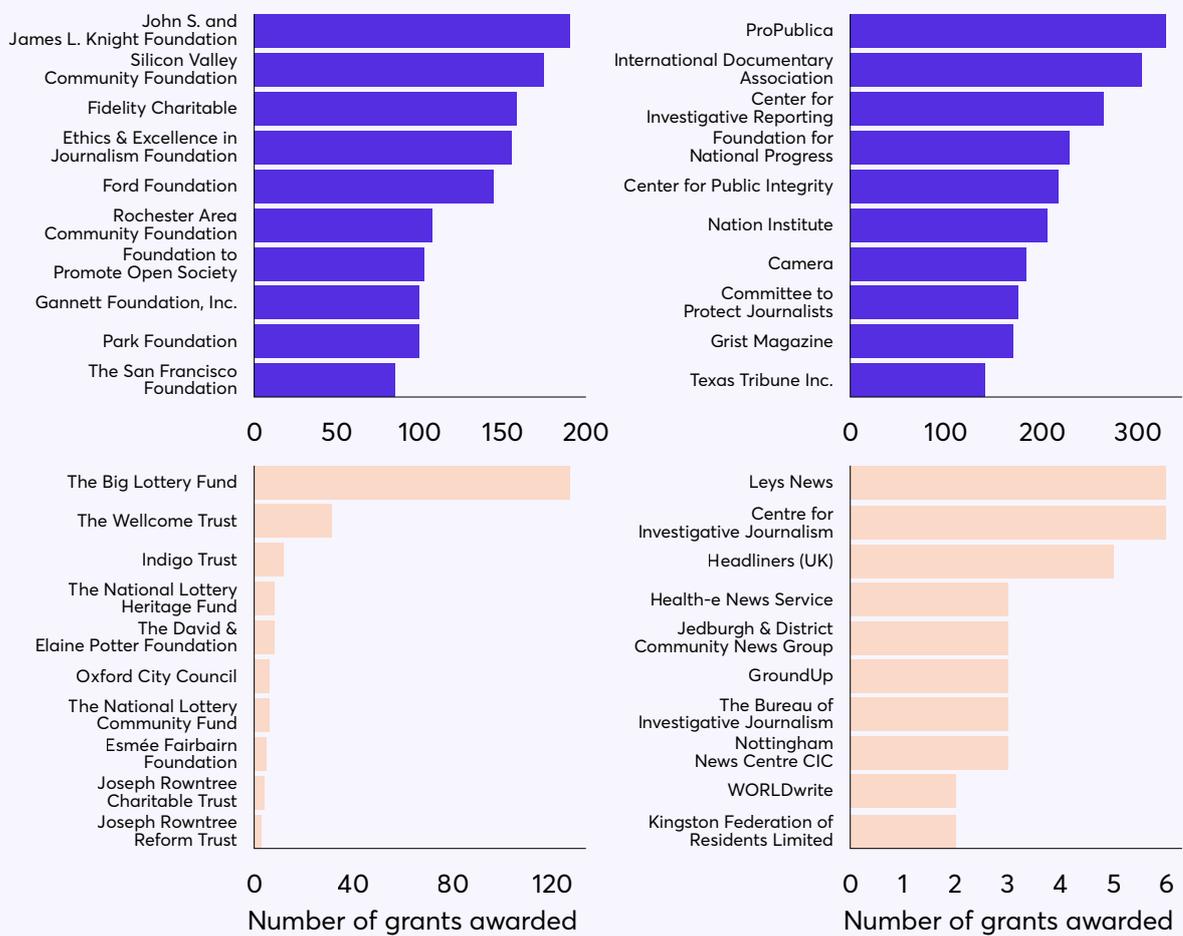
30 This is due to availability of the Candid data as explained in the Datasets section of report Appendix.

31 It should be noted that a limitation of the UK data used is that it does not include several important initiatives such as the BBC Local Democracy Reporting Service. The LDRS is not captured in the 360Giving dataset because it does not give out investments or grants, but instead awards contracts for which BBC retains the intellectual property.

(Figure 13). Despite a difference in the number of projects awarded per organisation (the maximum being 329 projects in the United States and six in the UK), it is clear that the UK coverage is quite low and not fully representative of the UK funding space.

Additionally, from the data available, the United States has a leading funder whose purpose is to invest in journalism - the John S. and James L. Knight Foundation.

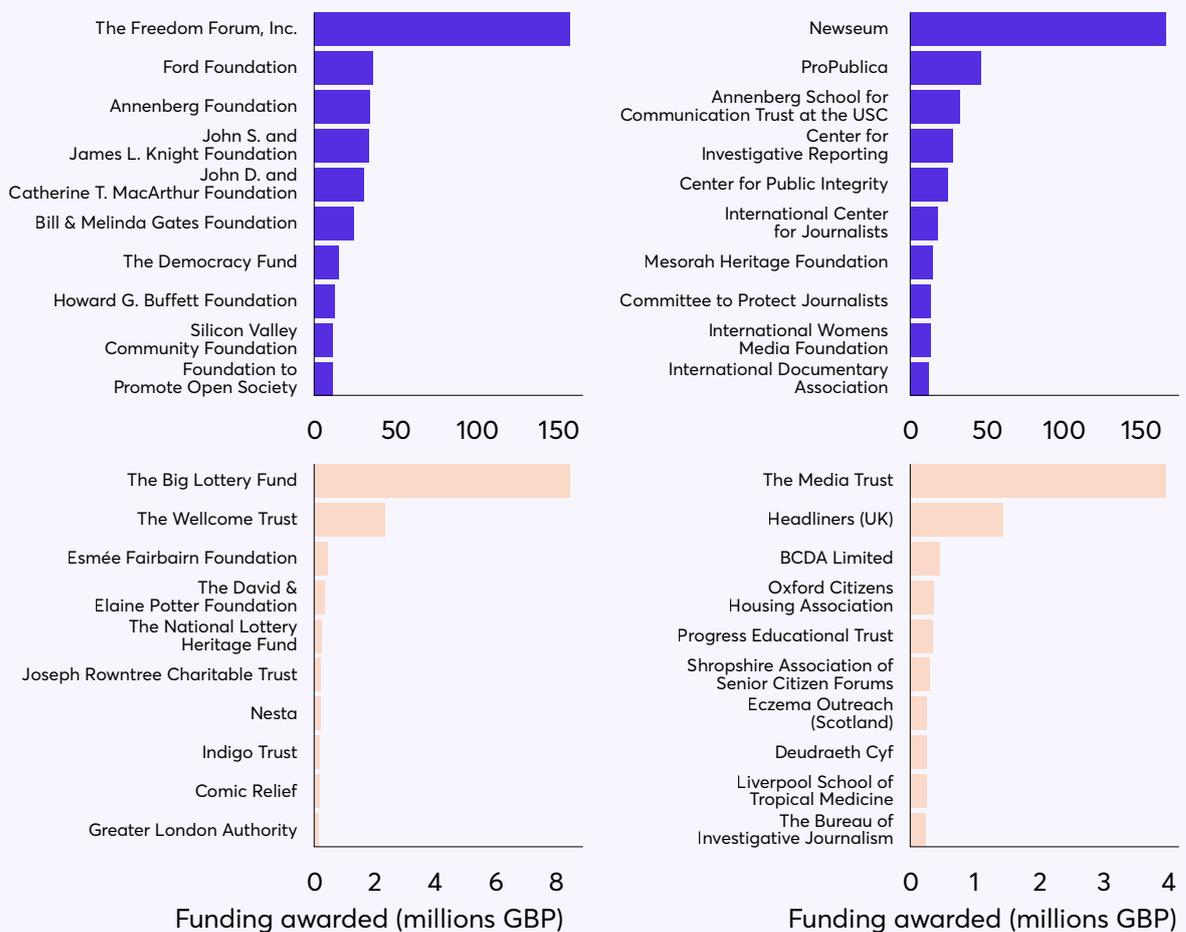
Figure 13: Top 10 US (green) and UK (blue) funders (left) and recipients (right) based on number of grants awarded or received



As shown in Figure 14, The Big Lottery Fund and The Wellcome Trust awarded approximately £8.5 million and £2.3 million of funding respectively, accounting for 79.6 per cent of UK spending on news and journalism grants between 2010–2019. The Wellcome Trust disbursed far fewer awards than the Big Lottery Fund, but awarded only slightly less in total funds, indicating that the projects it has funded are of higher average value. The Media Trust received the most funding (£3.95M) of all the recipients.

While funding concentration in the United States appears less concentrated than that of the UK, a large amount of funding was awarded by the United States organisation The Freedom Forum Inc.; awarding £156.5 million of their entire funding to the top recipient, Newseum (which received a total of £166 million, other funding included). No other funder awarded more than £47 million in the United States.<sup>32</sup>

Figure 14: Top 10 US (green) and UK (blue) funders (left) and recipients (right) based on total amount of funding awarded or received



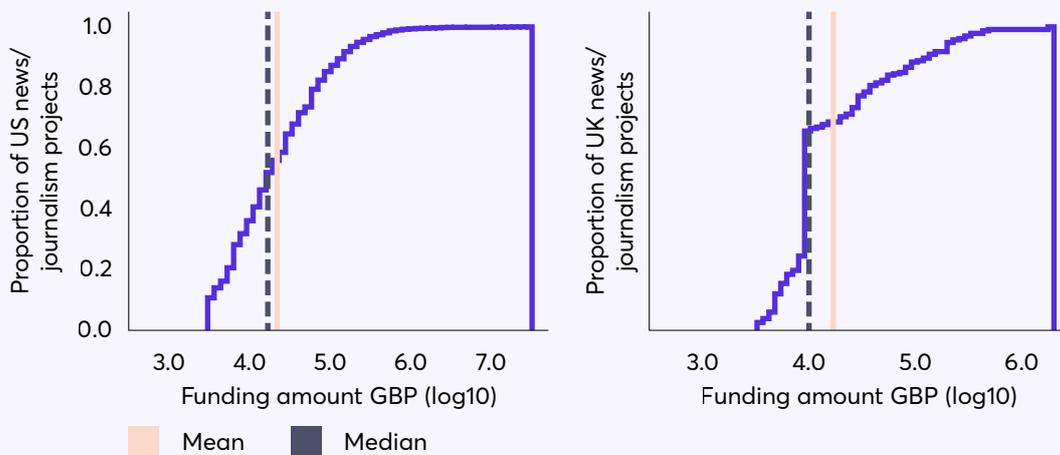
32 Again, it is important to note the discrepancies between the UK and US datasets. The low coverage of the UK may be due to a lack of data standard of a centralised data collector compared to the US. The US non-profit organisations complete 990 forms as a legal requirement which makes data collection for the U.S. much easier whereas non-profit organisations need to sign up to 360Giving to allow the data collectors to have access to their data. Since the sign-up process is not compulsory, many organisations may be left out or may not be aware of the sign up process.

**Funding distribution: UK vs. US**

Funding for news and journalism supports multiple types of activities and organisations. Funding can go directly to non-profit organisations to support improvements in journalism outputs, such as the US-based online magazine Grist. There is a large proportion of outreach and science communication projects, as well as research projects about methods such as fact-checking using machine learning and AI. There are also funds awarded to journalism training providers, such as the Centre for Investigative Journalism in the UK. Additionally, funding can be awarded to organisations whose aim is to support journalism organisations. One example is The Media Trust which has received funds from The Big Lottery Fund.

While there is a large difference between the total amounts of funding in news and journalism awarded and received by individual organisations in the United States and UK, the majority of projects received between £6,600 – £62,000. In Figure 15, we see that the mean and median in both the UK and United States are similar (in the £10,000s). In both countries, the mean (UK mean is £16,895, United States mean is £22,021) is greater than the median (UK median is £10,000, US median is £16,854) and the shape of the distributions suggest some projects being awarded very large amounts.

Figure 15: Normalised cumulative frequency distributions of funding amounts for projects in the US (left) and UK (right)

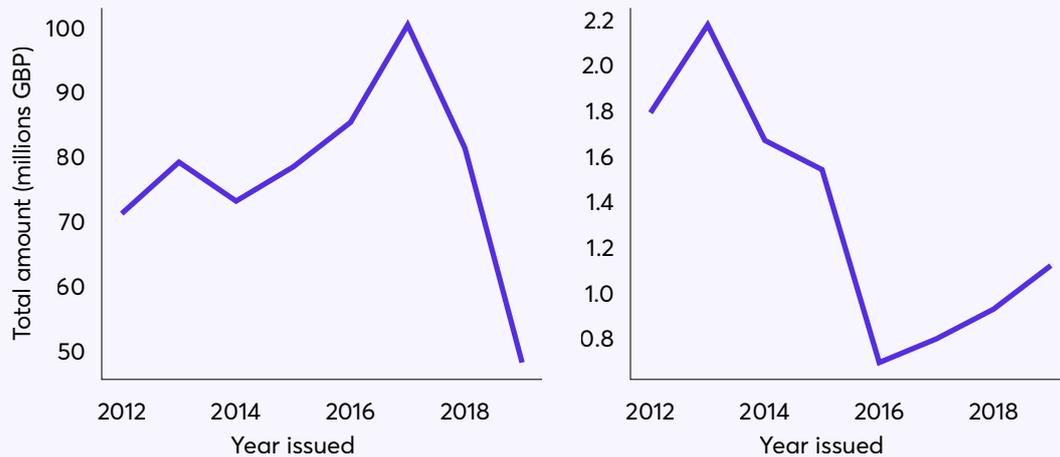


The analysis above aggregates funding and levels of funding activity from 2010– 2019. Additionally, we can look at the level of activity per year. Although there is a decline in the trend of funding for news and journalism, the number of projects is volatile over the years, staying fairly consistently at over 17 projects across all years but peaking in 2010 and 2015 (31 projects each), contrary to a low of 13 projects in 2011. Similar to the UK trend, the United States has a steady upward trend until 2017, with a slight dip in 2014.<sup>33</sup> However, there is a drop in the number of projects from 2017 onwards (starting at 1,110 to 28 projects in 2019) which is also reflected in the trend in spending (Figure 16).<sup>34</sup>

33 This upward trend from 2014-2017 may be an artifact of the data collection methods used by Candid, which relies on non-profit organisations filing for tax digitally (which is increasingly common).

34 This is due to a delay between the time when 990 tax form copies initially filed and when the copies are available for Candid to process the data.

Figure 16: Amount of funding awarded per year in the US (left) and the UK (right) (three year rolling mean)



### Sectoral funding comparisons

We now take a closer look at the UK to compare the level of activity between sectors. In order to make comparisons between sectors, we perform a word vector keyword expansion filtering (refer to the Methodologies section of this report). We used the project descriptions in the 360Giving data to find similar keywords to the sector names and used those keywords to filter on. The sectors chosen for this comparative analysis are:

- The sports sector (20,244 projects found)
- The health sector (28,406 projects found)
- The education sector (12,962 projects found)
- The technology and artificial intelligence (AI) sector (2,908 projects found)

Since we have pointed out the limitations of the 360Giving data, we cannot be confident to say that a sector has less projects missing and has been captured better as a result of having more projects selected during the keyword search compared to news and journalism. Hence, when we compare sectors for the analysis, we normalise the data to take into account the possibility of incomplete data for these sectors. Again, the data may not be completely representative of the national picture as it is difficult to capture.

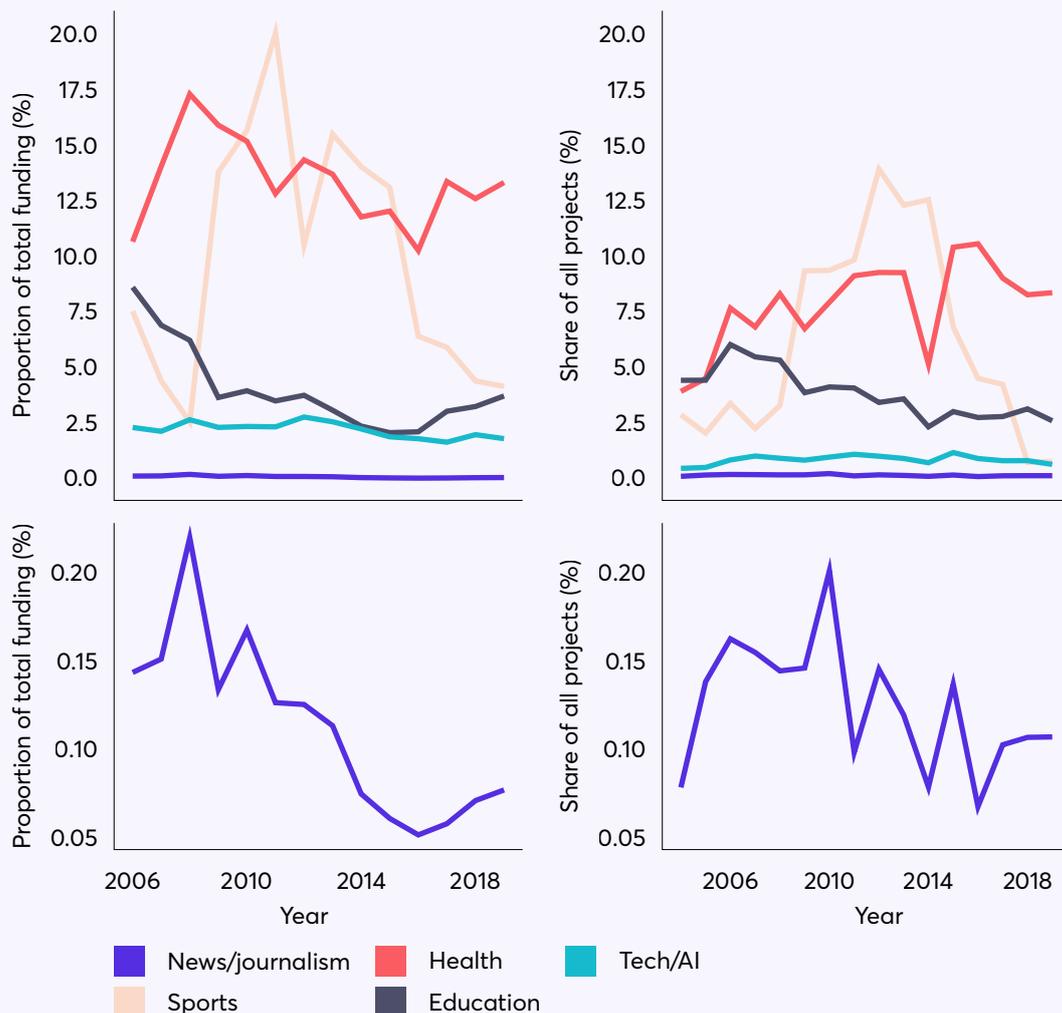
Between 2005–2019, the sports and health and education sectors individually made up a much larger proportion of total philanthropic projects than news and journalism, with sports peaking (between these sectors) at 13.9 per cent in 2012 (Figure 17). News and journalism projects consistently make up a very small portion of all projects peaking at 0.2 per cent in 2010 (Figure 17). This is similar in the technology and AI sector which makes up a share of less than 1 per cent across all years. A similar keyword search found that only 3 per cent of news and journalism projects are related to any of the technology and AI related themes.

Additionally, in Figure 17 again we see the total investment in news and journalism tends to have a small proportion (less than 1 per cent) of total funding across all years. In particular, news and journalism receive at most £3.1 million each year. We also see that there is a downward trend for the sector over the years, with the trough in 2016 being four times lower than its relative peak in 2008. This is unlike other sectors such as health where there is a slight

downward trend but the maximum relative drop is only 1.75 times lower than the peak in the trend. However in education, the trend of the proportion is also declining along with the trend of share of projects over the years. We note with interest that technology and AI projects had a similar share of all projects to news and journalism, however the proportion of total funding is consistently higher than in news and journalism (Figure 17). Although there are large levels of fluctuations in absolute terms, different sectors exhibit different trends.

As seen from the analysis of Crunchbase data presented in the previous section of this report, AI has recently received high levels of investment from the private sector as it is an emerging and cross-cutting technology. We may be starting to see a similar trend emerge in the non-profit sector as well, as the proportion of funding in technology and AI has started to increase from 2016 onwards. It is also interesting to see that the proportion of funding for health projects stays above 10 per cent across all years whilst this is not the case for the rest of the sectors at the end of the period.

Figure 17: Proportion of total projects per sector per year (graphs on left) and proportion of total amount of funding (graphs on right, three year rolling mean) for news and journalism projects and benchmark sectors



### Geographical spread of news and journalism funding and projects

In this part of the analysis, we explore the geographical spread of news and journalism projects in the UK.<sup>35</sup> Of the 431 UK news and journalism projects analysed in the 360Giving data, 332 projects have geographical information on the location of the beneficiaries. Due to the small number of projects, we decided to aggregate the original geographical information (which was at Local Authority Districts level) at NUTS 2 level.<sup>36</sup> All 41 NUTS2 regions have at least one project in news and journalism.

As seen in Figure 18, over 80 per cent of regions have at most 17 projects. However, Inner London - East has 45 projects (10.4 per cent of all news and journalism projects) and Manchester is second with 26 projects (6 per cent). It is also interesting to note that London as a whole, being composed of four regions, has 79 projects (18 per cent). In the same figure, projects in London, specifically Inner London - East and Inner London - West (with 17 projects), received the most funding in news and journalism (£8.5 million and £7.3 million, respectively). However, Northern Ireland is third receiving approximately £1.1 million in total. These findings highlight that even though a large percentage of activity may happen in metropolitan areas such as London and Manchester, spending is also distributed to non-metropolitan regions to a moderate degree. For example, two projects in Northern Ireland (recipients BCDA Limited and Headliners (UK)) alone received approximately £460,000 and £485,000, respectively. This is unlike the official data in Part 1a) where journalism employment hotspots are quite spread across metropolitan and non-metropolitan regions in the UK. Again, this may be due to the lack of full coverage of the UK in the 360Giving dataset.

Like the total amount of funding per region, the amount awarded for each news and journalism project (Figure 19) appears to be the highest in Inner London - West at £427,000, followed by projects in Inner London - East, receiving £182,000. Northern Ireland and South Yorkshire also each receive at least £72,300, surpassing the overall mean of £16,895. Despite the low number of projects per location, it appears as though non-profit investment is not concentrated uniquely in London. Taking into account the median (which is less impacted by outliers), on the other hand, shows that the amount of funding per project is much more widely distributed. Around half of the regions have a median amount ranging between £5,000 and £10,000, with a few regions having a median amount equal to the overall project median (£10,000) such as London regions (Inner West and Inner East), Kent and West Yorkshire. Many of the projects valued at exactly £10,000 were awarded by the Big Lottery Fund in these regions.

35 Again, this will be analysed within the limitations of the 360Giving dataset which doesn't fully capture the funding space in the UK but we hope will give a glimpse into where funding is going.

36 NUTS is the *Nomenclature of territorial units for statistics* used by the European Union.

Figure 18: Number of projects (left) and total amount of funding (right) per NUTS 2 region

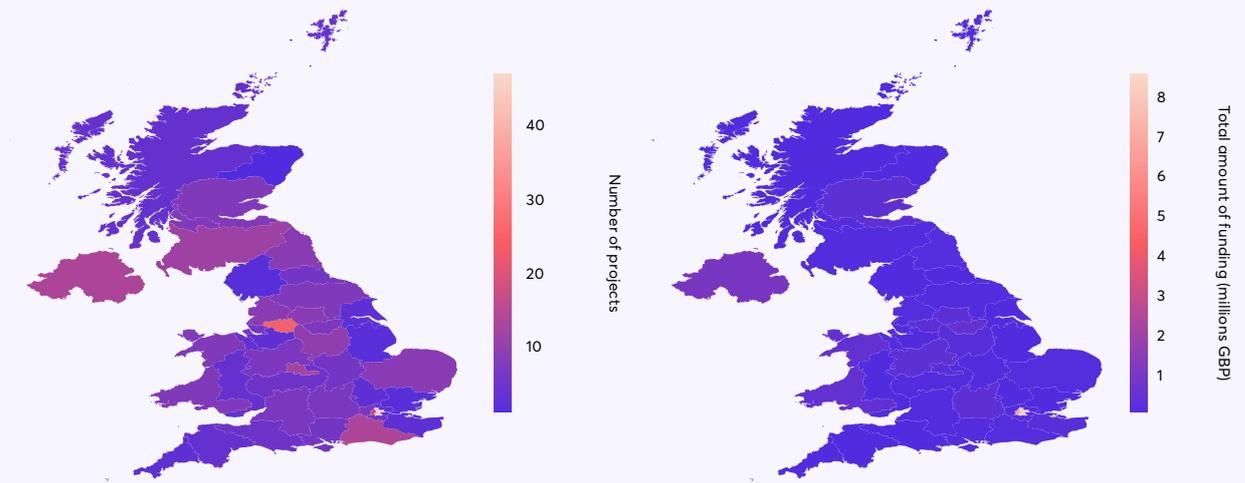
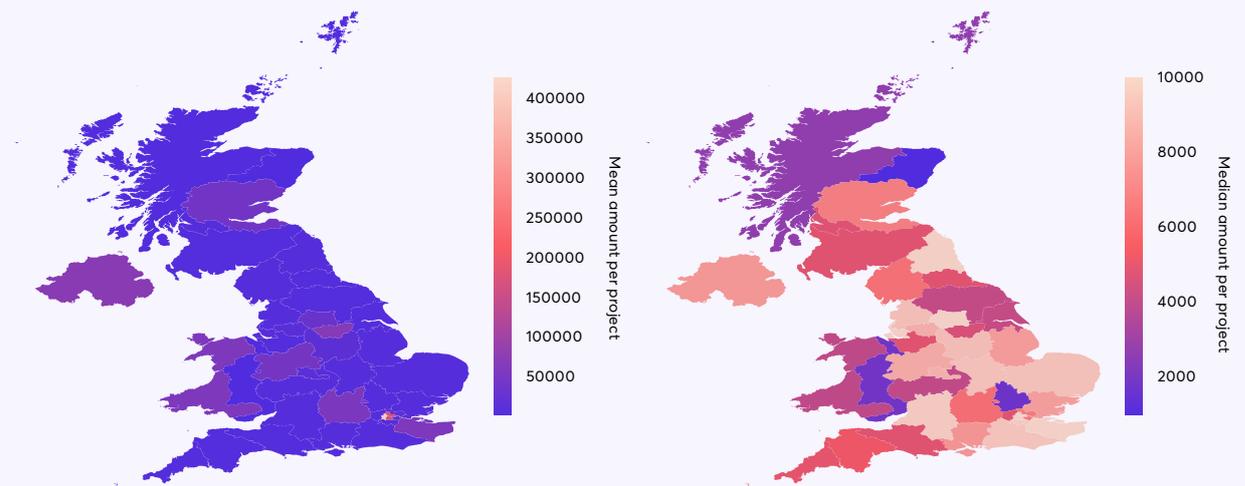


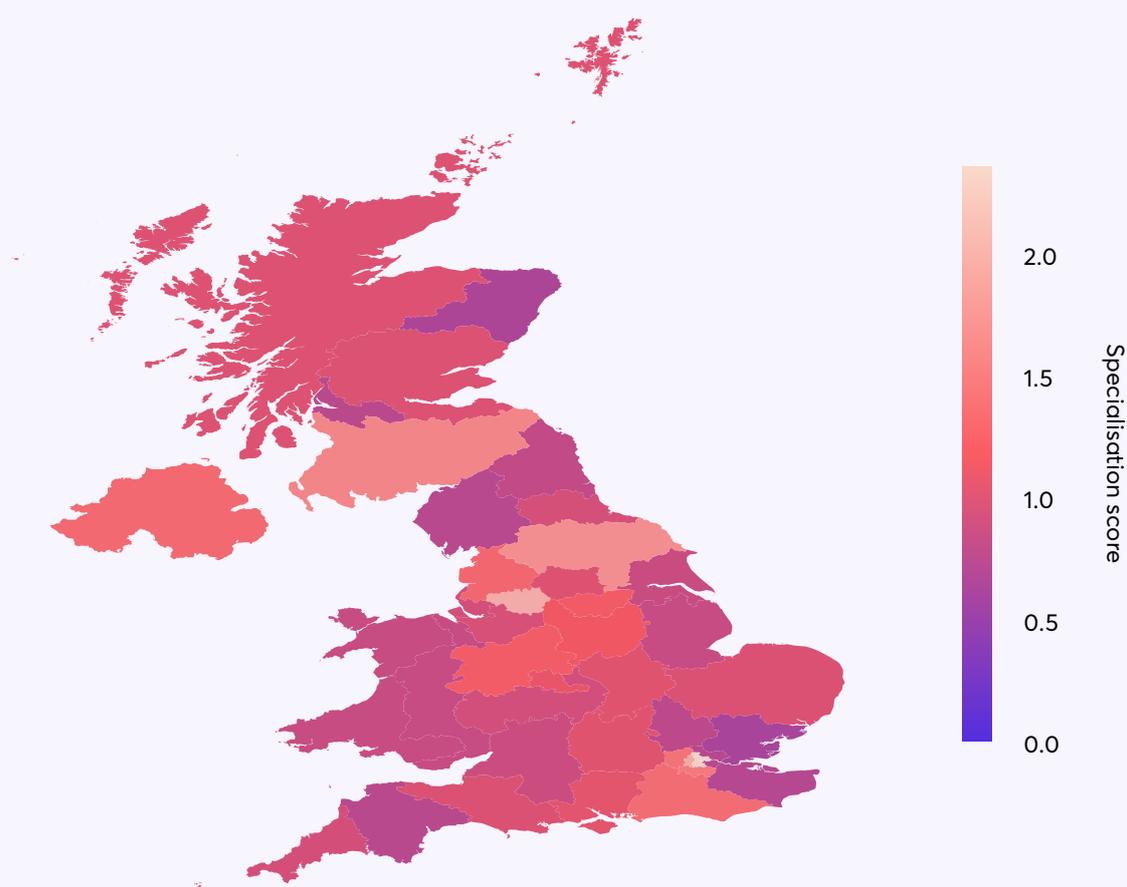
Figure 19: Log mean amount (left) and log median amount (right) per project per NUTS 2 region



Finally, we investigate whether there is a regional distribution of specialisation in news and journalism projects across the UK. In this part of the analysis, we use the geographical information of all projects from the 360Giving data (311,217 projects have geographical information). Again, we find that all NUTS 2 regions have at least one project.

In Figure 20, we see the majority of NUTS 2 regions are under-specialised in news and journalism (regions in purple). However, several regions have a degree of specialisation in this field (shown as the green regions: the higher specialisation, the darker green regions). Interestingly, these NUTS 2 regions are scattered across the UK. Two areas that are the most specialised, according to the data, are Inner London - East and Greater Manchester. Both have scores around two which means that there are two times as many news and journalism projects in these locations than we would expect given their overall presence in the 360Giving data.

Figure 20: Map of the specialisation scores per NUTS 2 region



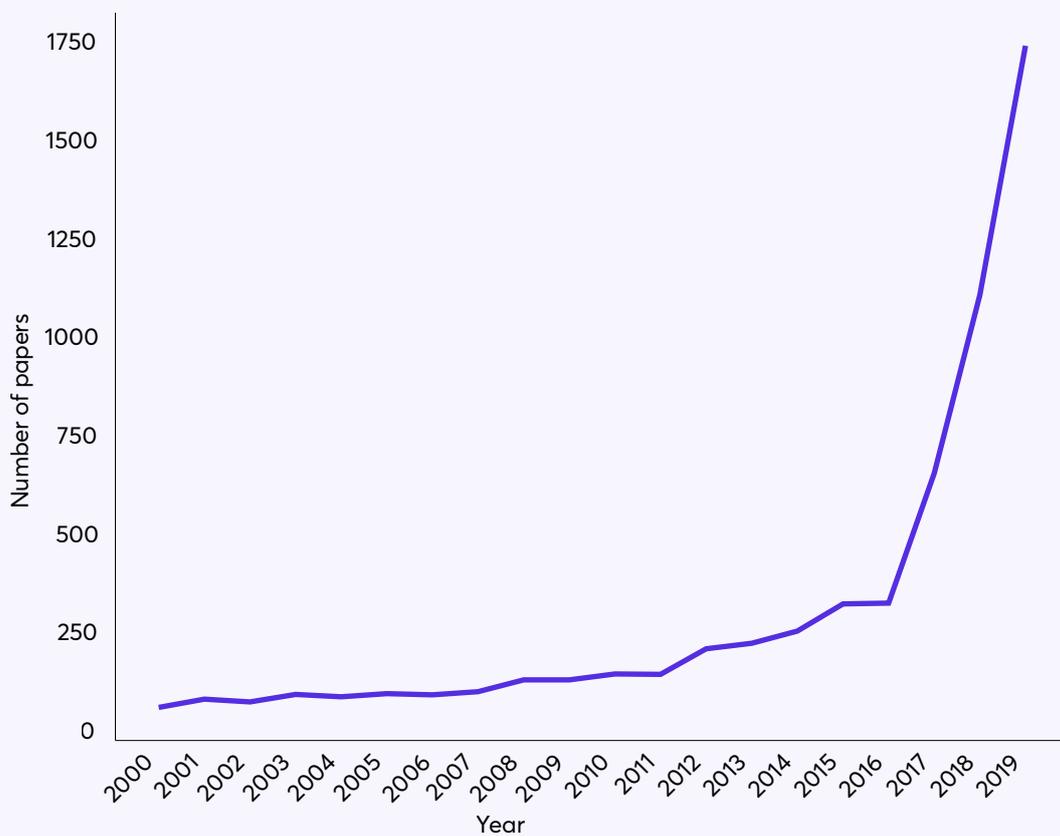
## Academic research activity on misinformation and disinformation

In this section we analyse trends, volume and geography of research activity on misinformation. We describe how research interests have changed over time, especially after the 2016 US elections. We identify the most cited researchers in the field as well as the institutions and countries with the highest volume of publications.

We analyse data from Microsoft Academic, a widely used open database with more than 240 million academic documents. Microsoft Academic uses machine learning to generate fields of study, a hierarchy of keywords that describe the academic publications. We queried it with keywords such as disinformation, misinformation, fake news and half-truth, and collected all of the papers that were published between 2000 and the end of 2019<sup>37</sup> that contained at least one of these keywords. We retrieved information on the paper, such as where it was published, publication date, number of citations, abstract and fields of study, and authors (names, affiliations). We then used the Google Places API to find the location of every author affiliation. Overall, we collected 13,081 papers, 21,213 authors and 1,932 unique affiliations, making this the largest mapping exercise in disinformation research that we are aware of.

Figure 21 shows the trend in misinformation research since 2000, showing a sharp increase in 2017.

Figure 21: Trend in published misinformation research (2000 – end-2019)



37 Later research could focus on the misinformation trends that have emerged since the start of the covid-19 crisis.

We examined the usage of fields of study (i.e. paper keywords) through time, splitting the analysis into Fields of Study (FoS). Figure 22 shows that the field of medicine has historically been the most used FoS in the dataset and has been consistently used throughout the time frame covered in the analysis, however it lost prominence in 2017 (following the US elections) when it was surpassed by fields such as computer science, social media, political science and media studies.

Figure 22: Fields of study in misinformation research (2000 – end-2019). Count of papers shown on y-axis.

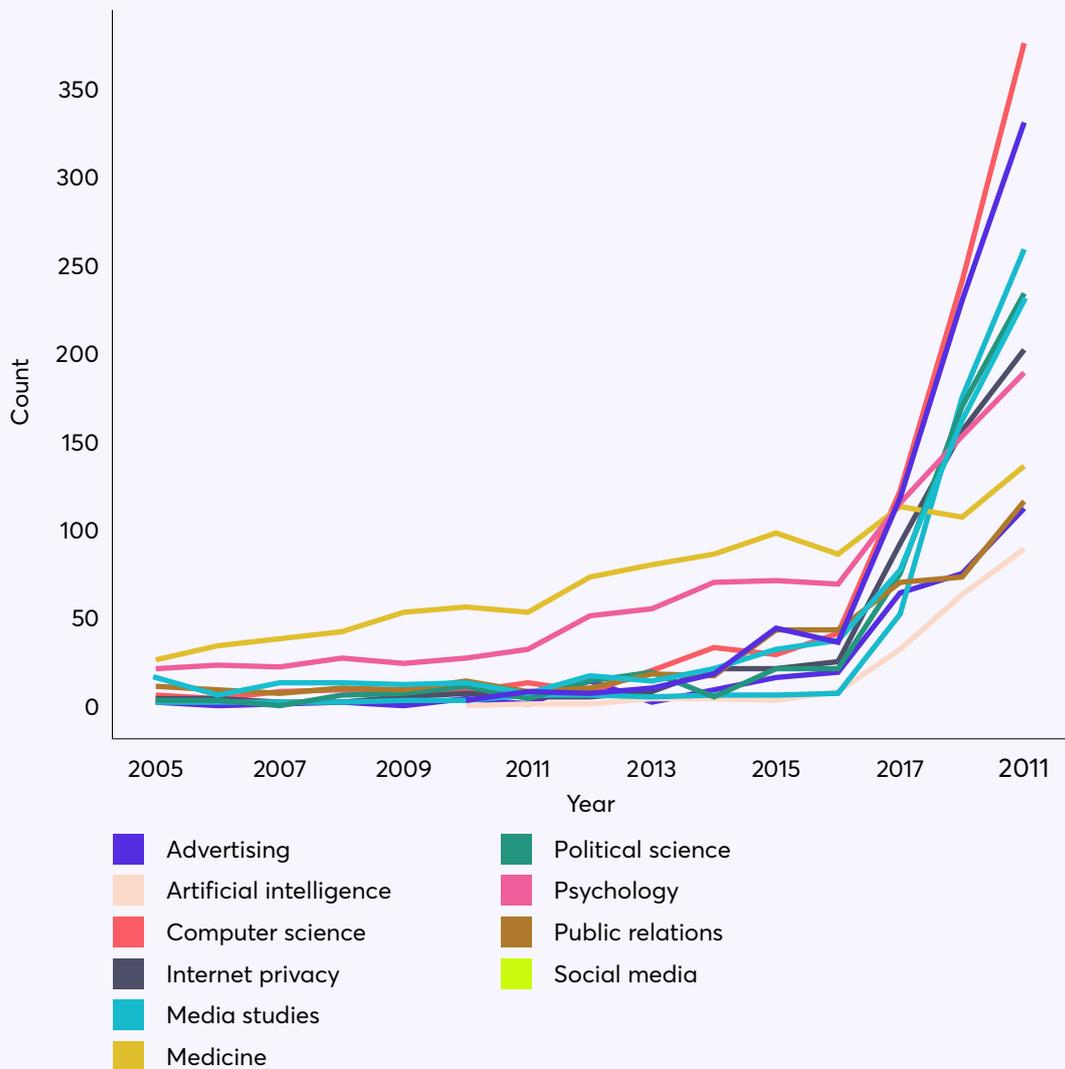


Figure 23 shows how the usage of terms related to misinformation has changed in Microsoft Academic. Misinformation is the term that has been historically used, however, fake news became the primary term to describe work in the field from 2017 onwards. Disinformation has the largest growth rate of all, even though it is still the least frequently used among the three terms. The term half-truth is rarely used.

Figure 23: Trends in misinformation-related terms in Microsoft Academic (2000 – end-2019). Count of papers shown on y-axis.

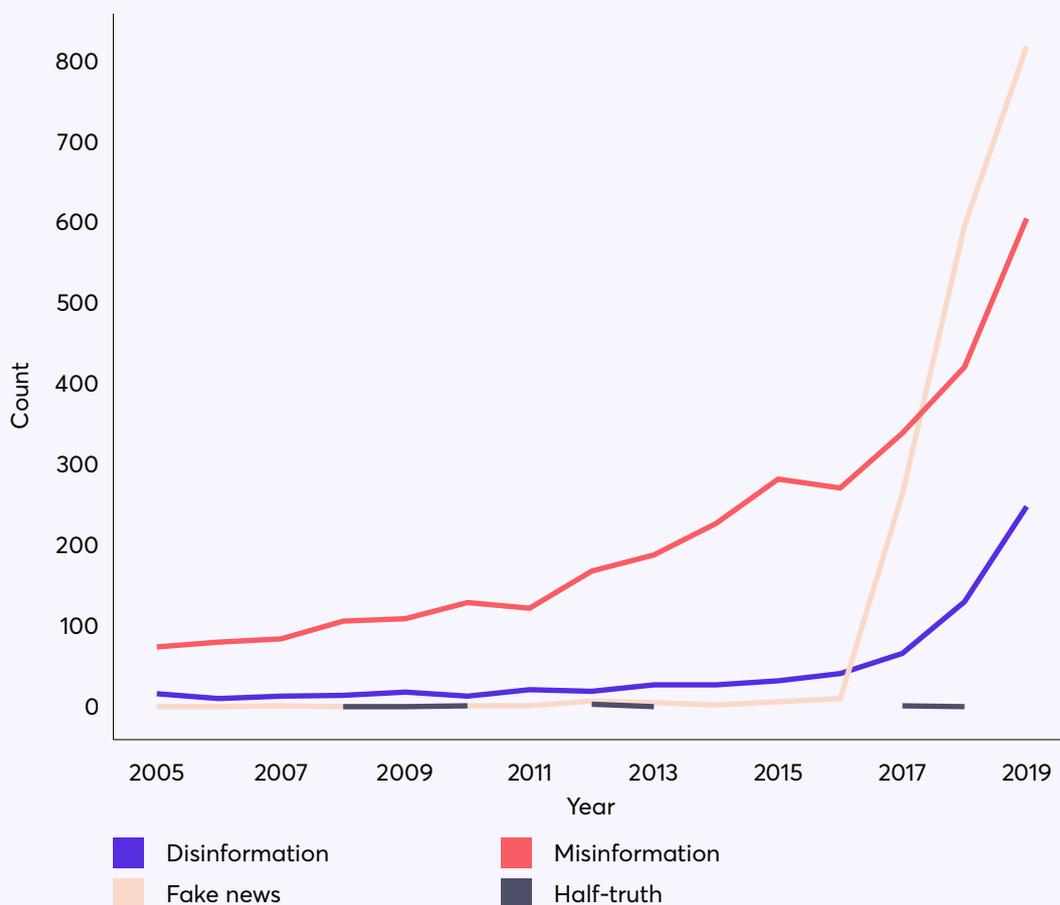
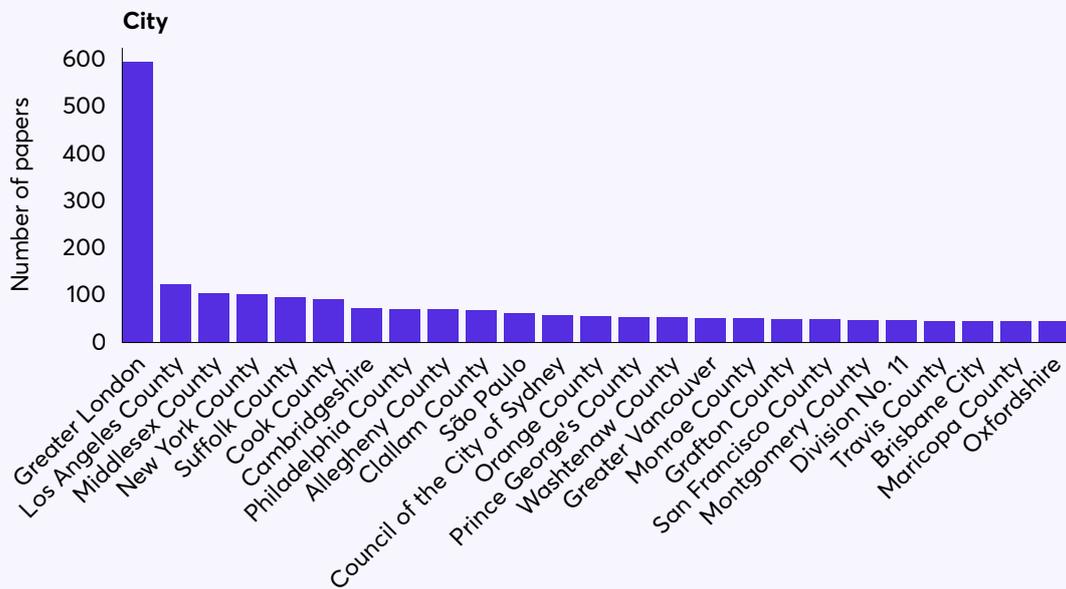
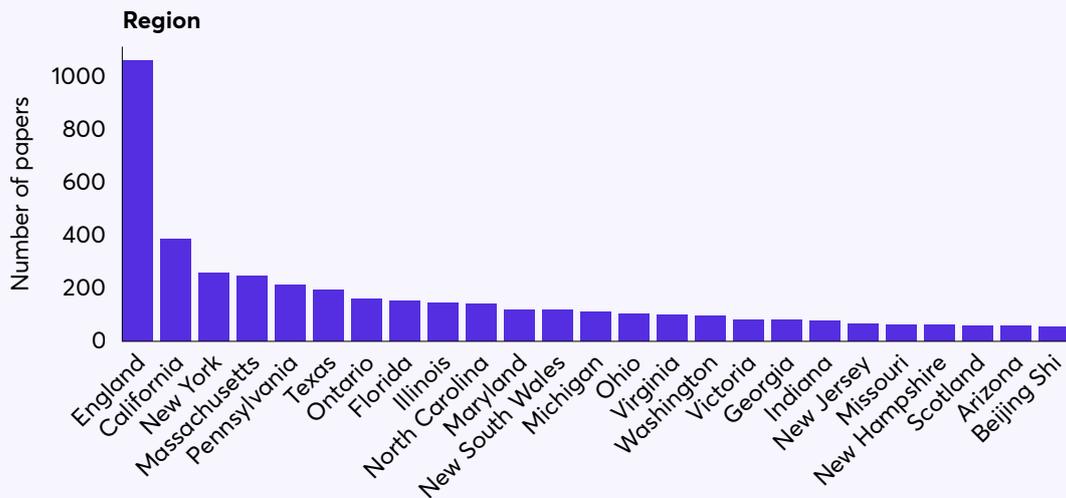
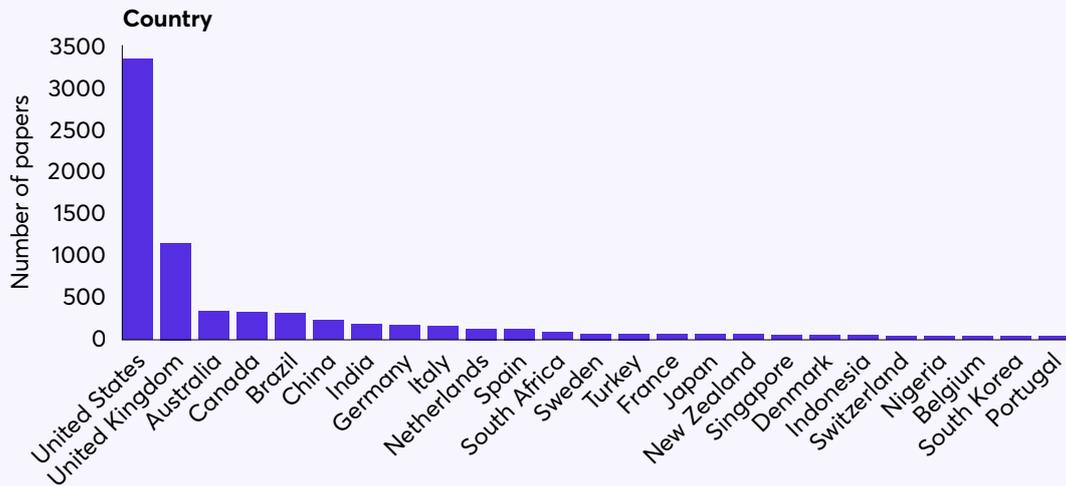


Figure 24 shows the distribution of misinformation-related papers between 2000 and end-2019 by country and region. The United States is the most active country in misinformation-related research, with a total of over 3,000 papers published in the study timeframe. The United States is followed by the United Kingdom (UK), with just over 1,000 papers, followed by Australia, Canada and Brazil.

Although the United States published the most research in this domain, the regional breakdown shows that this output is spread across several regions (e.g. California, New York, Massachusetts). By comparison, the output in the UK is almost entirely from institutions in England. Greater London has the highest output, followed by Los Angeles County (bottom of Fig. 24).

Figure 24: Distribution of misinformation-related papers (2000– end-2019) by country (top) and region (middle) and administrative area (bottom)



## 6.2 Innovation in public interest news

While Part 1 of this report focused largely on the sector news and journalism as a whole, the DCMS/Nesta Pilot Fund is aimed towards nurturing innovation in public interest news (PIN). No single definition of PIN exists, however the fund has supported a group of organisations that can be classified into several strands of news-related activities with public interest value. As a result, this part of the report looks at five dimensions of public interest news that are closely related to the activities of the Fund's grantees and investigates the level of company activity for each of them in Crunchbase. This aims to provide a nuanced understanding of public interest news and to highlight the gaps and opportunities within news and journalism sectors that have a high public value and where our grantees will operate. The sectors which have been investigated are inclusive news, legal and civic news, high-quality news, news revenue and news engagement, definitions of which can be found in the research questions and methodological summary (p.p. 15).

No categories for these sectors exist in Crunchbase, but the database does provide company descriptions, which we have used to generate a list of keywords that were subsequently used to tag companies involved in the five PIN fields. To generate the keywords, we began with lists of seed terms covering the themes. These lists were expanded to five larger lists of relevant terms using a bottom-up, data-driven method known as word vector keyword expansion. We used these keywords to perform a search for companies within the news, journalism and social news categories on Crunchbase, and after filtering out the 10 per cent least relevant companies according to their term frequency-inverse document frequency (Tf-Idf) ranking, we obtained five separate lists of PIN organisations.<sup>38</sup>

It is important to remember when interpreting the following analyses that our definitions of PIN do not necessarily require news organisations to be involved in journalistic activities, but can also include other organisations such as those providing services to the industry, analysing the industry or aggregating content from journalists. Additionally, while the low relevance threshold ensures that we capture as many relevant companies as possible based on their descriptions, it also increases the number of false positives. In this light, most figures relating to funding and levels of activity in the five public interest news sectors should be interpreted as over estimates.

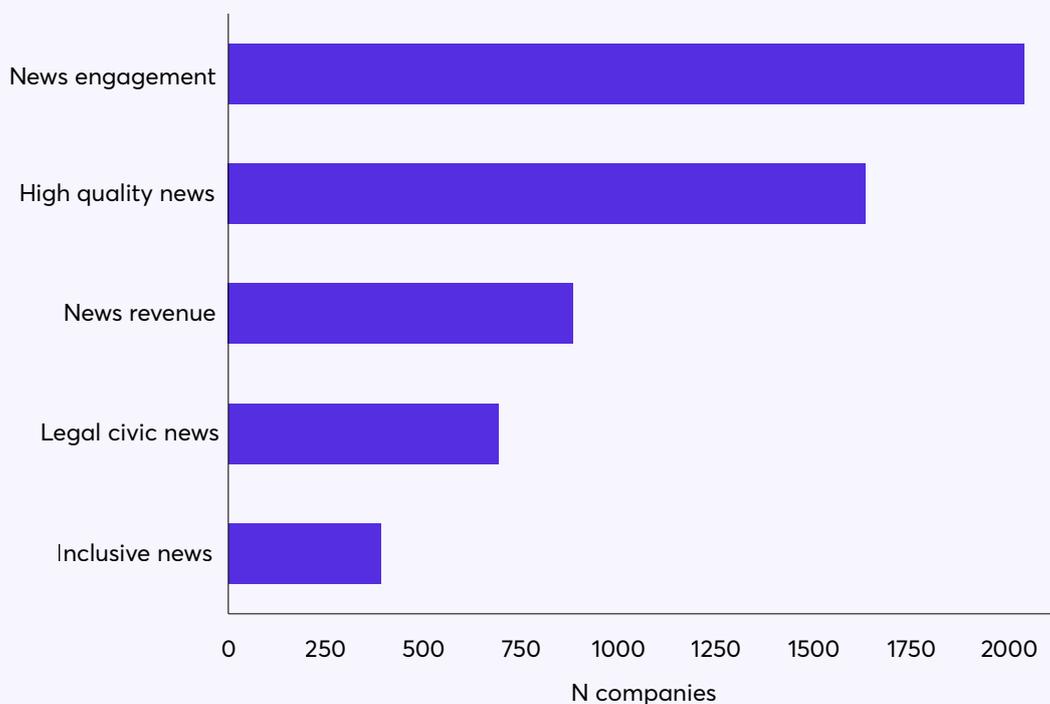
The following analysis investigates the levels of activity, funding landscape, geographic distribution and proliferation of technology within the PIN sectors.

### Public interest news companies globally

We identified just under 4,200 companies working across these public interest news themes. The breakdown between them can be seen in Figure 25, which shows that there are just over 2,000 companies working on engagement models, which include social media companies, forums, various forms of audience participation and engagement. This is closely followed by over 1,600 organisations that we have defined as part of the high-quality news sector, which are involved in investigative, factual and objective reporting. There are considerably fewer companies identified in the other categories; almost 900 for news revenue, which includes advertising and subscription services, and 700 for legal and civic news which covers reporting related to the legal and public sectors. Finally we identified just under 400 inclusive news companies which serve audiences including women, people of colour, the LGBTQ+ community, young people and more.

<sup>38</sup> Word vector keyword expansion and Tf-Idf ranking are explained in the Methodology section of the report and the full list of keywords is shown in the definitions section

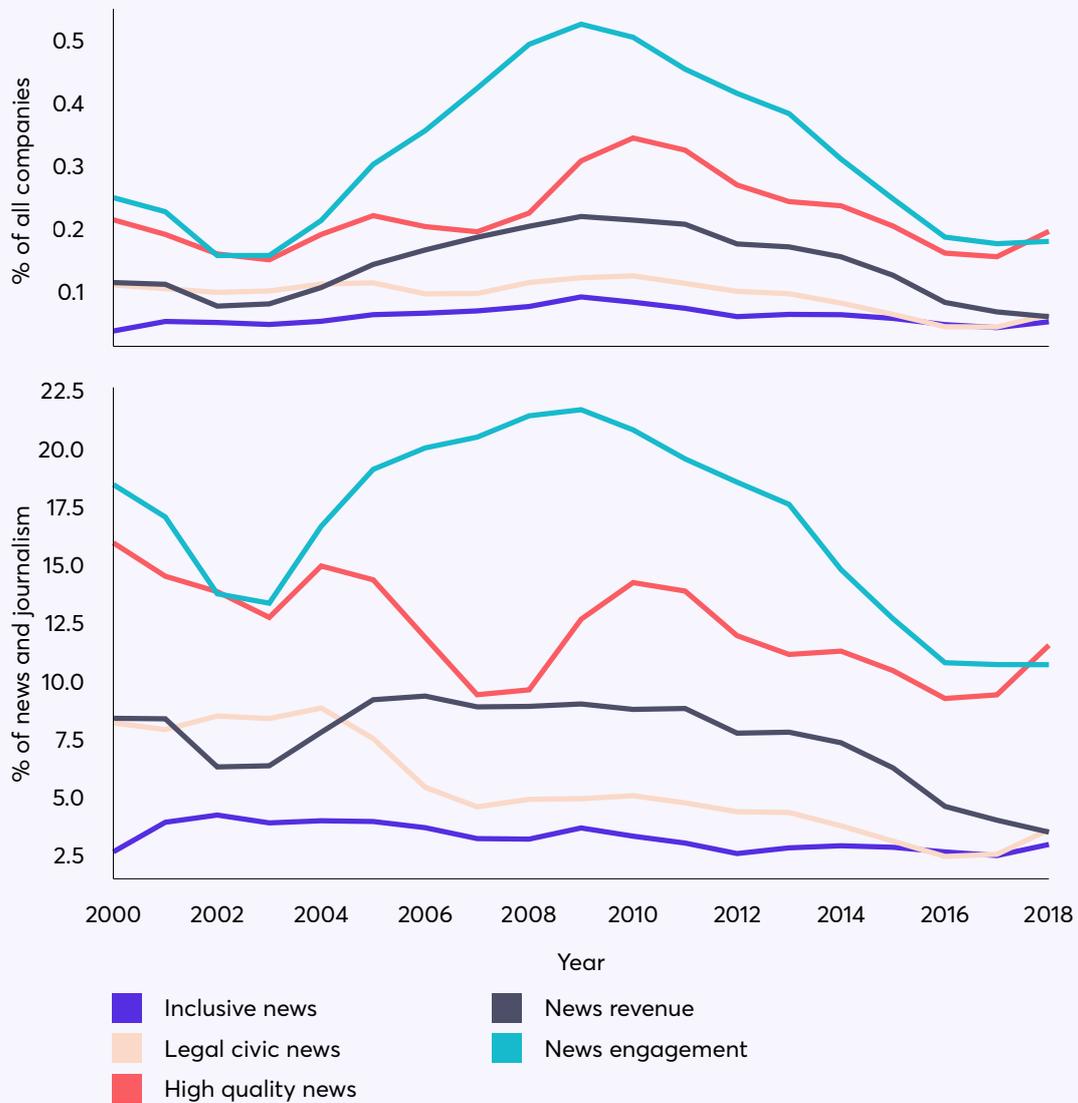
Figure 25: Total number of companies in Crunchbase across different themes of public interest news.



These five themes together appear to account for a significant portion of news and journalism companies in Crunchbase, but the levels of activity for each of them has followed a separate trend since 2000. Figure 26 shows the trends in these themes as a proportion of all company activity and as a share of news and journalism activity in Crunchbase. Similarly to news companies, relative activity in all of the PIN sectors peaked between 2008 and 2010. In particular, news engagement, which has a strong overlap with the social media sector, represented over 0.5 per cent of all companies in 2009, but this has since declined by almost two thirds. Between 2005–2018, news revenue rose significantly above legal and civic news to become the third most active PIN sector, but this has since declined and they have converged again.

Globally, all of the PIN sectors are experiencing similar or lower levels of activity as 20 years ago. It is not necessarily the competition between these themes that is important, but rather the overall trend as a share of all news and journalism activity. The lower plot in Figure 26, points to the decline in the importance of PIN across the news and journalism sector, supporting the idea that maximising clicks does not favour PIN and that other monetisation strategies are failing to attract investment. Knowing that both the absolute number of news companies has declined as well as the proportion of companies within the public interest themes, it is fair to say that these are being pushed out, with market failure being one potential explanation.

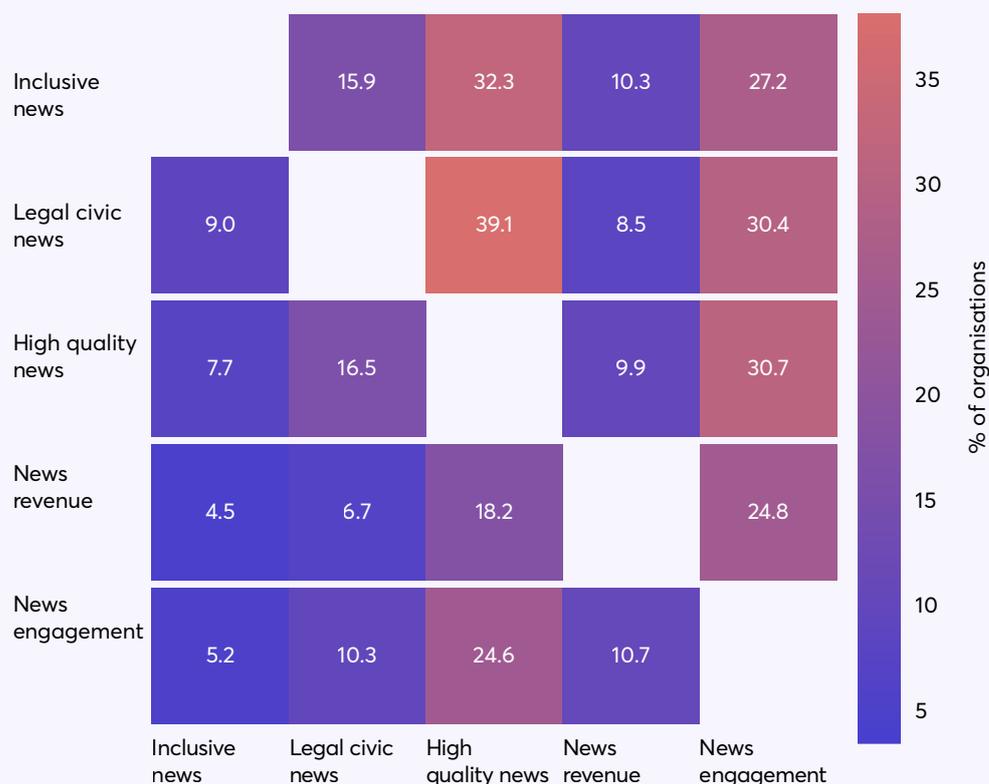
Figure 26: Level of activity over time in PIN thematic areas.



Naturally we see some overlap between themes, with some companies working across multiple sectors. While the vast majority (over 3,400) of companies working on these themes focus on just a single area, 28.3 per cent of them cover two or more of the themes. In Figure 27, we can see the proportion of companies within a public interest theme that also relate to each of the other themes. Each row of the heatmap shows the percentage of companies in a sector which have also been identified as belonging to another PIN sector. high-quality news and news engagement overlap the most with the other themes, perhaps unsurprisingly due to their higher overall frequency (as seen in Figure 27). There is a particularly high degree of crossover between high-quality news and legal and civic news, with 39.1 per cent of legal and civic news companies also being identified as high-quality news companies. While this is the largest overlap of any two categories, it also highlights that over 60 per cent of the descriptions of legal and civic news companies did not meet the search criteria for high-quality news, which includes terms relating to objectivity, factual reporting, investigative journalism and media bias.

Interestingly, all of the categories cross over with a lower proportion of news revenue companies than any other theme, despite news revenue being one of the smaller categories. This suggests disaggregation of these activities between companies, for example, those that are focused on generating new revenue models for news tend to be separate companies from those creating news content.

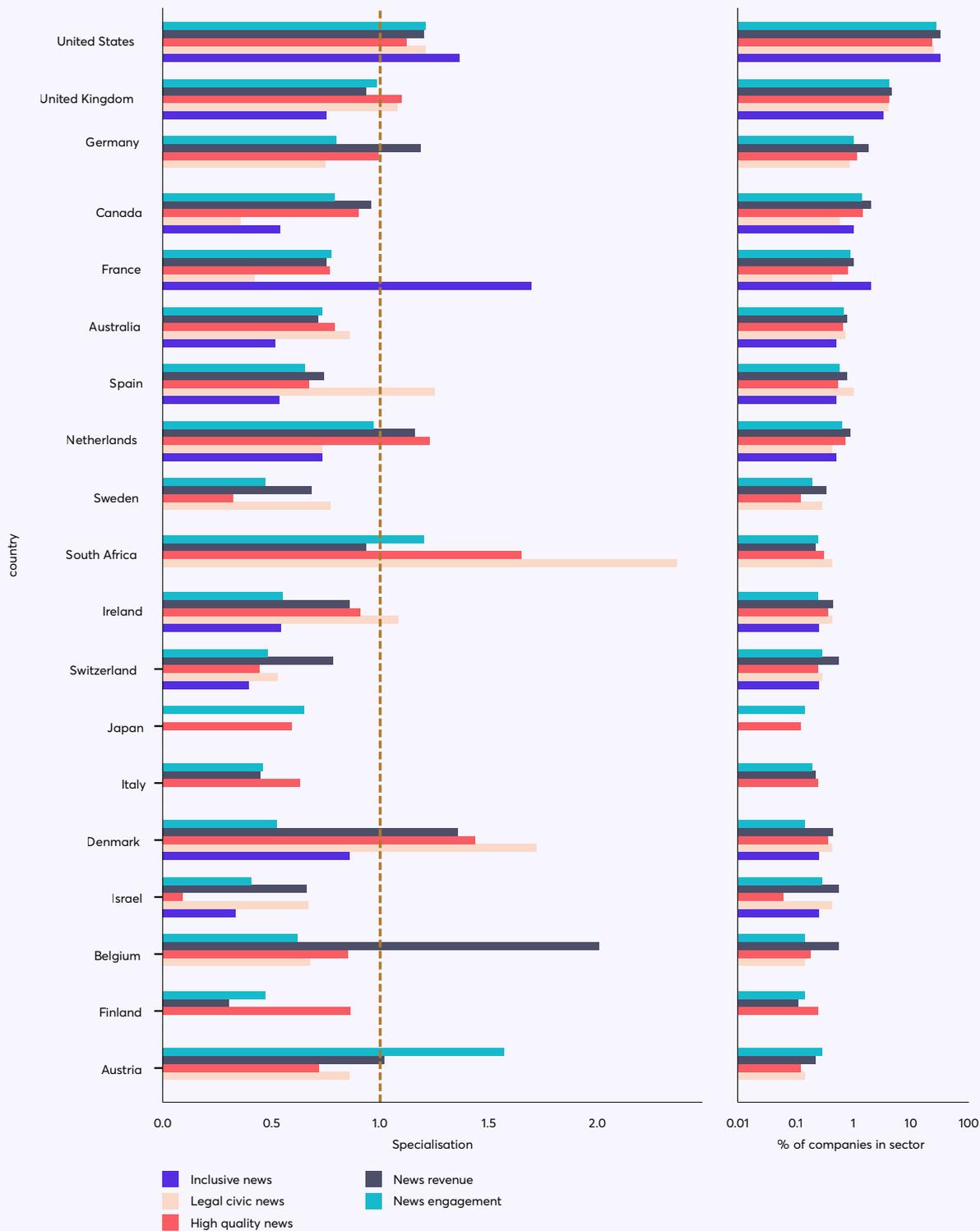
Figure 27: Level of crossover between PIN sectors



There is of course a geographical dimension to the distribution of activity in the public interest themes. Figure 28 shows that the United States is home to the largest share of companies for all of the five public interest themes by a significant margin, with the UK coming in second for every category. As discussed previously, we must bear in mind the limitations of Crunchbase when looking at raw counts of companies, hence we explore the degree of specialisation in each theme.

The United States is the only country that has a slight degree of specialisation for every PIN sector. It is joined by France as the only two countries with a significant specialisation in inclusive news. The UK has a varying profile of specialisation across the PIN sectors, being significantly under-specialised in inclusive news, and moderately under-specialised in news revenue and news engagement, while having a specialisation in high-quality news. Most of the other comparison countries exhibit under-specialisation across most of the PIN sectors. As these are generally high-income democracies, this raises questions about the drivers for the emergence of public interest news organisations. There are some notably high specialisations in particular sectors for certain countries, however this is due to measurement error arising from the low count of companies identified in these places. Very high specialisations should be interpreted with a high level of caution as they are likely to be much lower in reality.

Figure 28: Levels of activity for the top countries by number of companies and their specialisation in public interest themes. Note the logarithmic axis for activity levels.



## Emerging technologies in public interest news

Many of the industrial categories in Crunchbase represent technologies that can be applied within many other sectors. We identified categories that have experienced high growth rates in the last decade and may be applicable to news and journalism. These are AI, information technology, blockchain, the internet, mobile apps, big data and social media. In Figure 29, we can see the uptake of these technologies within public interest news companies in Crunchbase between 2002 and 2018. In almost all cases across this time span, there has been some level of growth in the uptake of these technologies among public interest news companies. However, figures and trends discussed here should be interpreted with caution due to the changes in fashionable terms used to describe emerging technologies and the fact that as technologies, such as the internet, become commonplace, companies are less likely to mention them explicitly in descriptions of their business activities.

In general, the internet is seen as playing a large role in both PIN and news companies, while the technology sector that has experienced the highest level of growth within the PIN sector is social media. In particular, the use of social media in news engagement has been particularly high, and has grown faster than for news companies, rising to almost 20 per cent penetration. Inclusive news has also seen a significant increase in overlap with social media, perhaps due to the uptake of social media platforms across wider demographics or an increasing use of social media to reach wider audiences.

There are other notable trends across the different technologies. In general PIN companies are less likely to overlap with the information technology sector than either news companies or the wider population of companies in Crunchbase. Perhaps surprisingly, there has been relatively slow crossover of companies across all sectors, besides news engagement, with mobile apps. This is beginning to change, although continued disaggregation is seen between mobile apps and news revenue. On the other hand, news revenue, along with legal and civic news and high-quality news have seen a more rapid crossover with the AI sector and a higher take up of big data than news companies in general. For news revenue this can be partly explained by the proliferation of machine learning and AI methods in serving advertisements and customising content, however it is less clear why this is the case for the other two sectors. For blockchain, rapid growth is visible in only the legal and civic news sector.



While we have discussed the relative proportions of technology that companies within news and public interest news are adopting, a more in depth analysis would be required to identify the exact nature of these companies across the globe. For example, a company classified under a public interest news sector and AI does not necessarily use the technology to carry out journalism, but may provide AI services to journalists or may be an organisation that reports stories from the AI sector. As examples within the scope of this report, we have extracted the names and excerpts of the descriptions of five companies identified as both within a public interest news sector and AI in the Crunchbase data which were founded in the UK. These are shown in Table 1.

Table 1: Names and descriptions of companies identified as public interest news and AI founded in the UK.

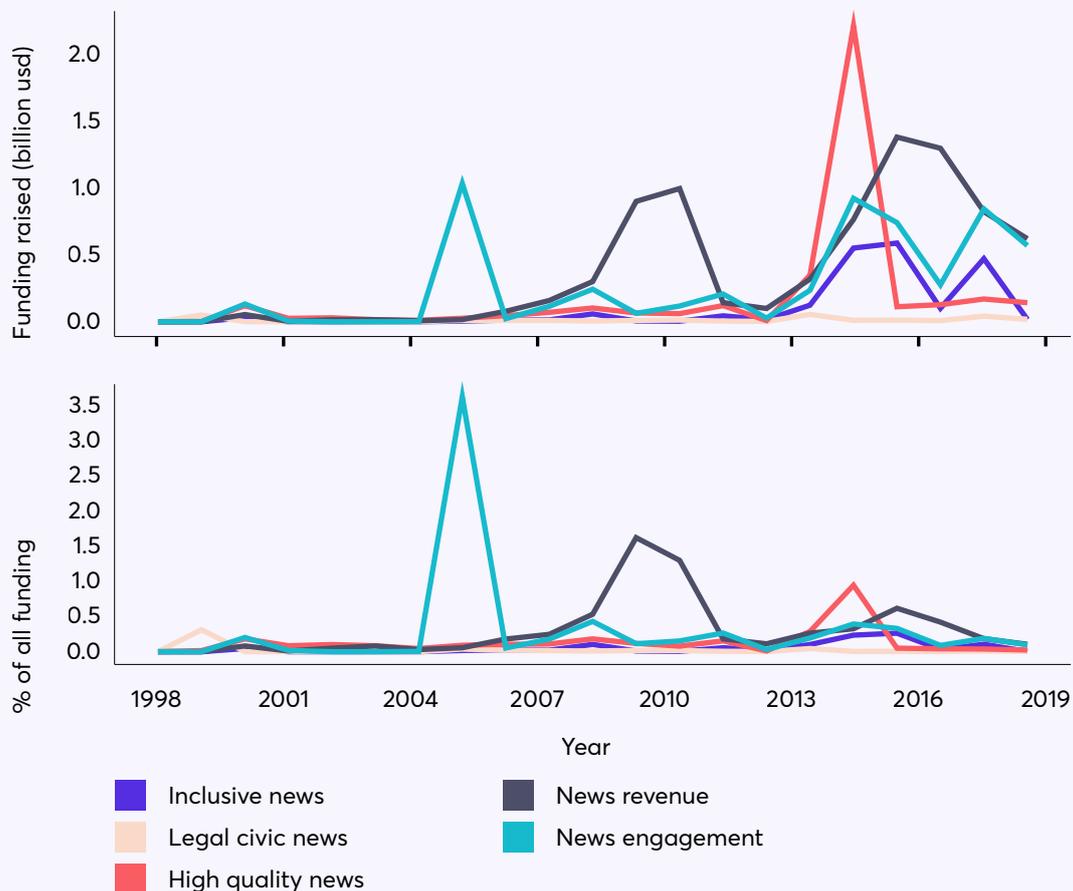
Company name	Description excerpt
<b>loomi.ai</b>	Loomi.ai offers a personal assistant capability to users to help them be more productive and achieve a better work-life harmony by managing automating information management. ... All information is also processed for factual accuracy and bias offering a personal assistant capability communicated to the user through online and voice channels.
<b>Krzana</b>	Krzana, the newsgathering engine for journalists, provides the tools to surface pre-news and uncover the narrative behind local and global events, before they become news. Our platform's deep linguistic analysis supports fast, trustworthy journalism in some of the world's most advanced newsrooms by alerting professional storytellers to breaking stories, and providing unbiased information from any location, in any language. ...
<b>Echobox</b>	We are building the first AI that understands the meaning of content. Our product helps publishers automate their social media management and grow their audience. ... Echobox Insights provides analysis and advice on social media, AI and journalism in the digital age.
<b>Missum</b>	Missum is an intelligence service that delivers stories that interest you, straight to your mobile phone. Its unique twist is its ability to detect and unveil the varying sides of every story, so that you can obtain a spherical view. ... A content recommendation service for inquisitive minds!
<b>Logically</b>	We have been described as a "fake news search engine" and "the credibility layer to the internet." In a nutshell, at Logically we use Natural Language Processing, machine learning and human oversight to identify bias and misinformation on any given topic. News is assigned a credibility score and ranked on the site accordingly with credible information appearing up top. ...

## Private sector funding for innovation in public interest news

As well as varying profiles of technology overlap, the PIN sectors we explore have different funding profiles, both in terms of the amount of money raised by companies and the sources of those funds. In this section of the report, we analyse the level of investment across the five themes and what types of funds have been used to support PIN both in the UK and the 18 other comparison countries.

Globally, it appears that funding in most PIN sectors has increased over the last two decades. Figure 30 shows that four out of five of the sectors have raised between \$100 million and \$2 billion in any given year, however legal and civic news significantly lags behind in terms of growth and volume. Also notable is the fluctuating rise and fall in absolute investment, with all of the sectors having experienced one or more funding bubbles. These are in fact caused by large, individual instances of funding for a small number of organisations, for example the 2005 peak in news engagement is due to a \$1 billion corporate fund from Google to AOL Lifestream. The more recent peak in the inclusive news sector is a result of six funds and investments to Vice Media with values between \$100 million and \$450 million. By looking at the relative share of funding for all PIN sectors in terms of overall funding listed on Crunchbase we see that the underlying trend beneath these anomalies suggests public interest news continues to attract a very minimal share of funding.

Figure 30: Levels of funding in PIN sectors between 2002 and 2018.

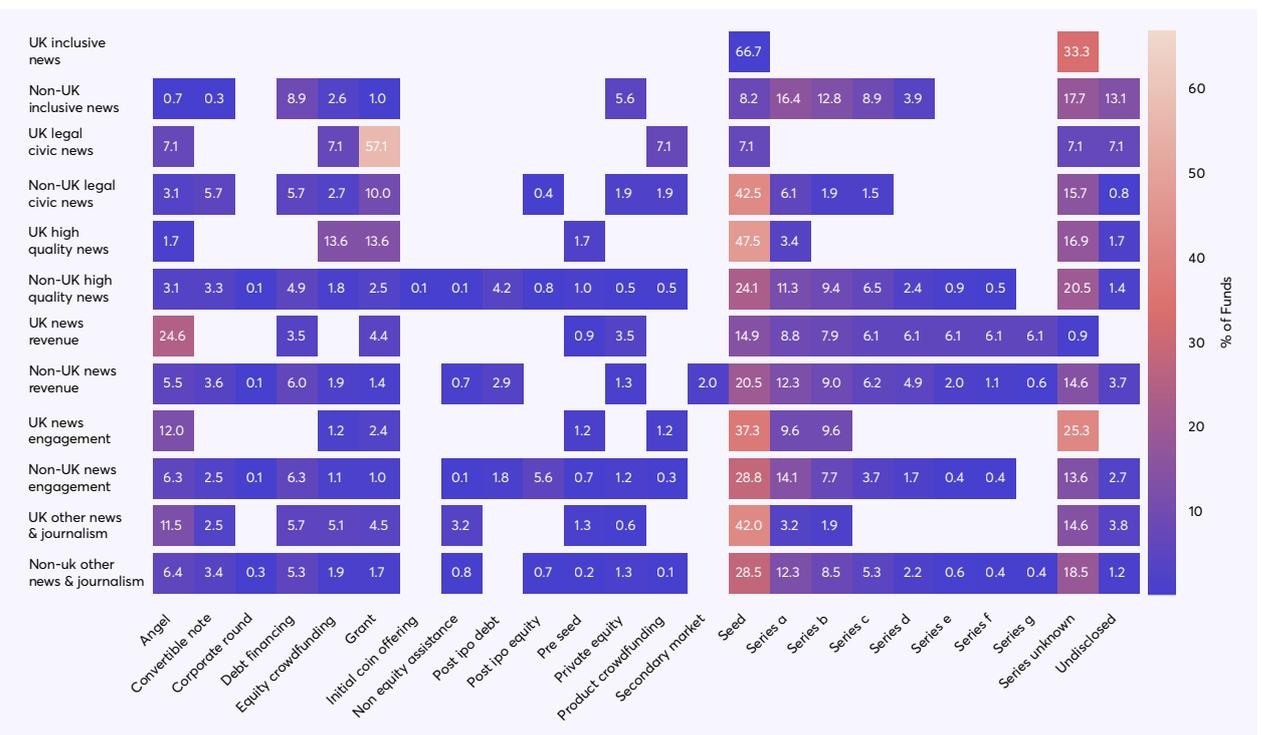


There are of course multiple ways that an investor can choose from to make funding available for companies or that a company can choose to raise funds, including direct investments, grants or crowdfunding. We see that the pattern is different across PIN sectors and between the UK and the 18 comparison countries (although we note that most activity is accounted for by the United States). In the UK, funding for news and journalism in general is not too dissimilar to the global picture, with notable reliance on seed, venture capital, angel investment and grant funding. This changes significantly when we look at the individual PIN sectors.

For inclusive news, companies we observe a strong presence of seed and venture capital, which is similar to the global pattern (although it should be noted that this relies on only five instances of funding recorded in the UK). In general, across the four other PIN sectors, we see a stronger reliance on angel investment, grants, product crowdfunding, equity crowdfunding and seed investment although the exact profile and the degree to which this is the case differs between sectors. Apart from news engagement and news revenue, we see a much lower presence of venture capital in the UK compared to the United States and other countries.

While fund amounts are not available for all of the funding instances in Crunchbase, we also show the median levels of funding for the various fund types across PIN sectors and news and journalism. The average funding in the UK is lower across every sector, but this is unsurprising as Section 1b) of this report highlighted. There are some exceptions. The average grant value is somewhat higher in the UK than for legal and civic news and news revenue companies, while the average angel investment for high-quality news in the UK is also higher than the global comparison average. Finally, for news engagement in the UK, the first three rounds of venture capital are also higher than the average in the comparison countries. This suggests again that the UK suffers from a patchy investment and funding landscape, but has certain avenues that are significantly more worthwhile for organisations to pursue.

Fig 31: Share of funding for public interest news in the UK and comparison countries by type of funding



# 7.0

## Conclusions and next steps

---

**This pilot study has demonstrated that it is possible to bring together multiple data sources on news and journalism activity, and to dig deeper into the public interest news landscape using new data sources and methods.**

Our analyses of journalism and media-related employment in the UK support the notion that there is an ongoing structural change occurring in the industry. While newspaper publishing remains the top employer in absolute numbers, it has progressively been losing ground to employment in domains such as web portals. Turning our attention to new company formation using the Crunchbase dataset, we find a decrease in activity following a peak in 2009. Despite some growth in funding for these ventures, investment in new enterprises appears to be quite low (particularly relative to activity and investment in companies focused on emerging technologies such as AI and blockchain).

We also explore the philanthropic funding landscape in the UK and United States using data from 360Giving and Candid, respectively. Bearing in mind the limitations of these datasets (described in detail in the body of the report), we find that philanthropic grants for news and journalism initiatives in the UK are on average smaller than in the United States, and that the funder landscape in the UK is far more concentrated with 73 per cent of funds coming from the top three funders (relative to 7 per cent in the United States).

Given the increasing urgency to understand and address misinformation and disinformation in news, we also explored the academic trends in this domain. We found a stark discontinuity between pre- and post-2016 research trends, with the former period being dominated by research related to medicine, and the latter characterised by politics, news and social media, internet privacy and fake news.

Using Crunchbase data and natural language processing methods, we have also been able to identify companies working in one of five sectors relevant to public interest news: inclusive news, high-quality news, legal and civic news, news engagement and news revenue. We have built a preliminary global picture of private and public sector activity in these sectors. While the data is biased towards the United States, we have cautiously extracted some global and UK level patterns in terms of founding activity, funding, technology penetration and sectoral overlap within public interest news.

Primarily, we find that public interest news sectors have occupied a decreasing share of companies being founded worldwide over the last decade. Within news and journalism, this trend holds true within the news and journalism sector, where the proportion of public interest news companies is slowly declining. Over the same time period, the share of funding registered in Crunchbase for the public interest news sectors has remained low at 0.2 per cent or below, although this is occasionally punctuated by individual investments with values over \$100 million.

There are some signs of modernisation in public interest news. We see evidence of increasing crossover with emerging technology sectors such as AI. Indeed, the growth and proportion of companies operating across public interest news and emerging technology sectors sometimes exceeds those values for Crunchbase as a whole. In particular we see strong recent overlap between social media and news engagement as well as AI with high-quality news, legal and civic news, and news revenue.

Finally, activity recorded in Crunchbase suggests that the UK has a mixed profile of specialisation across the public interest news sectors we have investigated. In particular, we see under-representation of inclusive news provision in the UK, and to a lesser degree some under-specialisation in news revenue. However, the country has a higher level of activity in high-quality news than we might expect. Overall, the UK performs better in terms of specialisation than many other countries with a similar political and economic profile. However, innovation in the sector may be limited due to the lack of a diverse funding landscape. The UK is reliant on grants, seed funding, equity crowdfunding and angel investments to support public interest news ventures, whereas venture capital, corporate funding and debt financing is a less prominent feature of the funding landscape.

The study has also revealed several important gaps in the data used to analyse the news and journalism globally, but in the UK specifically as well. These are outlined in detail in the limitations section, however it is clear that in order to promote a more cohesive and integrated understanding of the UK situation, more work is required to collect and make available robust data.

There are many aspects of this work that could be further developed by the Nesta team that conducted the research, or by others in the broader research ecosystem. For instance, the majority of the findings presented in this report are descriptive in nature, and while they shed light on many important questions, they raise a great deal of 'so what?' questions. In other words, there are also many questions that can be explored about what these trends mean to outcomes related to democracy and the strength of institutions. We take a first step toward this kind of analysis when we explore the association between journalism employment and health outcomes using the Public Health England profiles. In future, these types of analyses could be expanded to other outcomes such as the Oxford Consultants for Social Inclusion (OCSI) Indices of Deprivation.

The geographic scope of the analysis could also be broadened, with more exploration into the activities and funding of UK organisations working internationally (e.g. Internews, Article 19, Index). Future research could also delve deeper into what innovations in public interest news globally have proven successful, and to explore whether they might be adapted for a UK context.

Finally, there are many additional data sources that could be explored in future. For instance, we explored the possibility of scraping news websites to better understand the extent to which their coverage focused on local issues. Performing primary data collection of this nature proved to be too time-intensive for the current study, however a number of services exist that are collecting the content of news so in future it could be possible to license this data or to obtain it through a partnership.

We welcome any comments, feedback or references to further work that would help to strengthen or challenge the findings presented here. Please get in touch with the research team if you'd like to start a conversation.

---

# Acknowledgements

---

This report benefited directly and indirectly from many people who were involved in the Future News Pilot Fund, including members of the Nesta UK programme delivery team, advisory board members, and many members of the wider media and journalism community who engaged with the research team throughout the project. Special thanks go to Sameer Padania and Isabel Newman for their invaluable feedback on a draft version of the report.

Carrying out the data analysis for this report would not have been possible without the availability of high quality open source and free-to-use quantitative computing tools. This report made use of many libraries and packages from the Python scientific computing ecosystem, particularly Pandas, Numpy, Matplotlib, Scikit Learn, Spacy and Gensim. We thank all of the maintainers, contributors and supporters of these projects and others.

# References

---

- Adshead, Stephen, Grant Forsyth, Sam Wood, and Laura Wilkinson. 2019. "Online Advertising in the UK." Plum Consulting. <https://plumconsulting.co.uk/online-advertising-in-the-uk/>.
- Cairncross, Frances. 2019. "The Cairncross Review: A Sustainable Future for Journalism." [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/779882/021919\\_DCMS\\_Cairncross\\_Review\\_.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/779882/021919_DCMS_Cairncross_Review_.pdf).
- DCMS. 2019. "Jeremy Wright's Statement on the Cairncross Review." GOV.UK. 2019. <https://www.gov.uk/government/speeches/jeremy-wrights-statement-on-the-cairncross-review>.
- Howells, Rachel. 2015. "Journey to the Centre of a News Black Hole: Examining the Democratic Deficit in a Town with No Newspaper." Cardiff School of Journalism, Media and Cultural Studies. <http://orca.cf.ac.uk/87313/1/2016howellsrphd.pdf>.
- Kirby, Dr. Philip. 2016. "Leading People 2016: The Educational Backgrounds of the UK Professional Elite." [https://dera.ioe.ac.uk/30232/1/Leading-People\\_Feb16-1.pdf](https://dera.ioe.ac.uk/30232/1/Leading-People_Feb16-1.pdf).
- Mediatique. 2018. "Overview of Recent Dynamics in the UK Press Market." [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/720400/180621\\_Mediatique\\_-\\_Overview\\_of\\_recent\\_dynamics\\_in\\_the\\_UK\\_press\\_market\\_-\\_Report\\_for\\_DCMS.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/720400/180621_Mediatique_-_Overview_of_recent_dynamics_in_the_UK_press_market_-_Report_for_DCMS.pdf).
- Newman, Nic. 2019. "Reuters Institute Digital News Report 2019," 156.
- Ofcom. 2019. "Review of BBC News and Current Affairs." [https://www.ofcom.org.uk/\\_data/assets/pdf\\_file/0027/139905/bbc-news-review-terms-of-reference.pdf](https://www.ofcom.org.uk/_data/assets/pdf_file/0027/139905/bbc-news-review-terms-of-reference.pdf).
- Padania, Sameer. 2018. "An Introduction to Funding Journalism and Media." ARIADNE: European Funders for Social Change and Human Rights. <https://www.ariadne-network.eu/wp-content/uploads/2015/03/An-Introduction-to-Funding-Journalism-and-Media.pdf>.
- Spilsbury, Mark. 2018. "Journalists at Work." Independent research commissioned by the National Council for the Training of Journalists.
- Wardle, Claire. 2018. "Information Disorder: The Essential Glossary." Harvard Kennedy School.

# Appendix: Data, definitions, methods and limitations

---

## Datasets

### 360Giving

360Giving is an open data standard endorsed by the UK Government, and aims to aggregate standardised information about the UK philanthropic funding space.<sup>39</sup> We collect the entire dataset (360,348 grants) from the open database and use word vector keyword expansion to create sector-specific data subsets. Data is available from the years 1991 to 2020. There are 126 unique funding bodies present in the data and 86.5 per cent of projects have geographical information available which provides full coverage of the UK Local Authority Districts (LADs).

### Business Register Employment Survey (BRES)

BRES is the official source of employee and employment estimates by detailed geography and industry in the UK, in the public or private sector and full time or part time basis. The data come from a survey of approximately 80,000 businesses, and is weighted up to represent the economy covering all sectors. More details on BRES data can be found here.

### Candid

Candid is a data aggregator of the global funding space. The collection process by Candid is done via information from non-profit tax filings (990 forms) of US organisations and sparse collection of non-US projects. Within the licensed dataset, Candid has categorised each project with one or more categories. The data we queried fall under "Information and communications" which has a selection of subcategories. Taking this into account, there are 136,542 projects, with 135,602 projects that are both US-based and US-funded. Due to the robustness of the collection of US projects, we will use the US-funded and US-based projects to compare to the UK data provided by 360Giving. Data from 2010 to 2020 is available in the data. Additionally, foreign exchange rates from the 1 of January of each year are used to convert the currency in the data (US Dollars) to Great British Pound Sterling (GBP) to compare.

39 See: <https://www.threesixtygiving.org/2019/03/18/government-sets-the-standard-with-360giving/> and <https://www.threesixtygiving.org/support/standard/>

### **Crunchbase**

Crunchbase is a paid, licensed global database of private and public sector companies. It includes details on companies, such as their founding date and country, and a description, as well as information about funds, investors and company founders. In addition, each company is tagged with one or more of 750 industrial sector categories, allowing for sector based analyses.

The database is regularly updated, and this report has used a snapshot of the full database containing information up to July 2019, detailing 783,806 companies. Crunchbase collects data through a combination of automated methods, an in-house data team, a community of online contributors and partnerships with venture firms, accelerators and incubators. While the database covers the majority of countries, the US accounts for one third of companies in the database, indicating some bias. In this report, we have attempted to balance the reporting of absolute numbers of companies and funds with relative measures.

### **Public Health England**

We use the Local Authority Health Profiles published by Public Health England to explore associations between levels of journalism-related activity and indicators of health and wellbeing. The Health Profiles programme is part of the Chief Knowledge Officer's Directorate of Public Health England. Data are available from 2006 to 2019, and the online tool is updated on a quarterly basis. The data are open and can be accessed via the following link: <https://fingertips.phe.org.uk/profile/health-profiles/>

### **Microsoft Academic**

Microsoft Academic is a widely used, open database with more than 240 million academic documents. It contains scientific publication records, citation relationships between those publications, as well as authors, institutions, journals, conferences, and fields of study. MA uses machine learning to generate Fields of Study - a hierarchy of keywords that describe academic publications, which can then be queried for analyses. The database is currently updated on a weekly basis. For more information, visit: <https://www.microsoft.com/en-us/research/project/microsoft-academic-graph/>

## Definitions

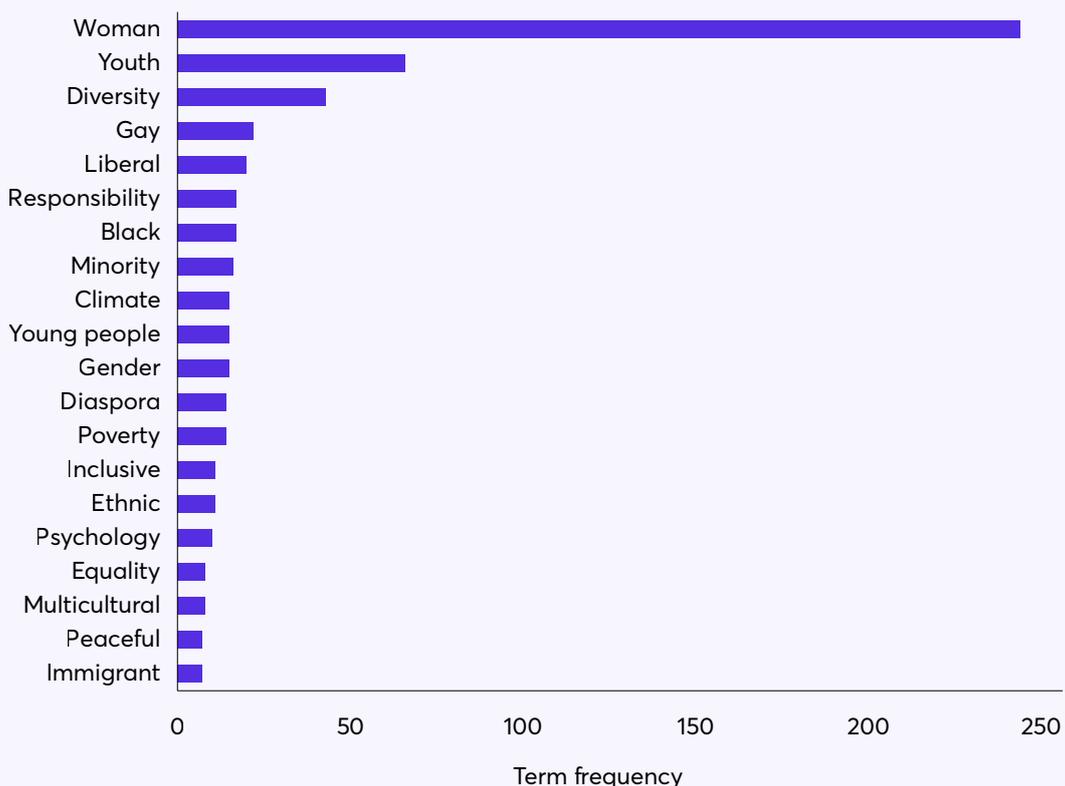
### Crunchbase

This section contains lists of keywords generated through word vector keyword expansion (see Methodologies) that have been used to search for companies relating to public interest news and themes related to the Pilot Fund within Crunchbase. Each set of terms is accompanied by a graph of the top 20 term frequencies across all of the company descriptions retrieved from Crunchbase using the search terms.

### Public interest news

abuse, affair, amendment, analysis\_commentary, analysis\_opinion, balanced, bias, biased, bureaus, citizen, city\_county, civic, civil, commentary, commentary\_analysis, comprehensive\_coverage, conservative, correspondent, corruption, courage, court, criticism, current\_affair, defend, democracy, democratic, depth\_coverage, detect, differently, dignity, disagree, discourse, diversity, election, equality, equitable, ethic, factual, fairness, fake\_news, federal\_government, federal\_state, freedom, governmental, hard\_hit, high\_standard, human\_interest, human\_right, ignite, ignore, immigration, incisive, inclusive, individuality, injustice, insightful, inspiring\_story, integrity, investigation, investigative, investigative\_journalism, investigative\_reporting, journalism, journalistic, journalistic\_integrity, justice, labor, law\_firm, lawyer, legal, libertarian, liberty, local\_government, mainstream\_medium, misinformation, misuse, moral, national\_security, negative, objectivity, openness, original\_reporting, participatory, peace, peaceful, policymaker, politic, politic\_crime, politic\_economy, political, political\_party, politically, politics, positive\_change, positivity, prejudice, propaganda, protest, provocative, racism, reform, relentlessly, remain\_committed, reporting\_analysis, reveal, shed, shine\_light, socio, stance, standards, state\_government, stimulate, stimulating, struggle, thought\_provoke, thoughtful, truth, underrepresented, underserved, uphold, violence

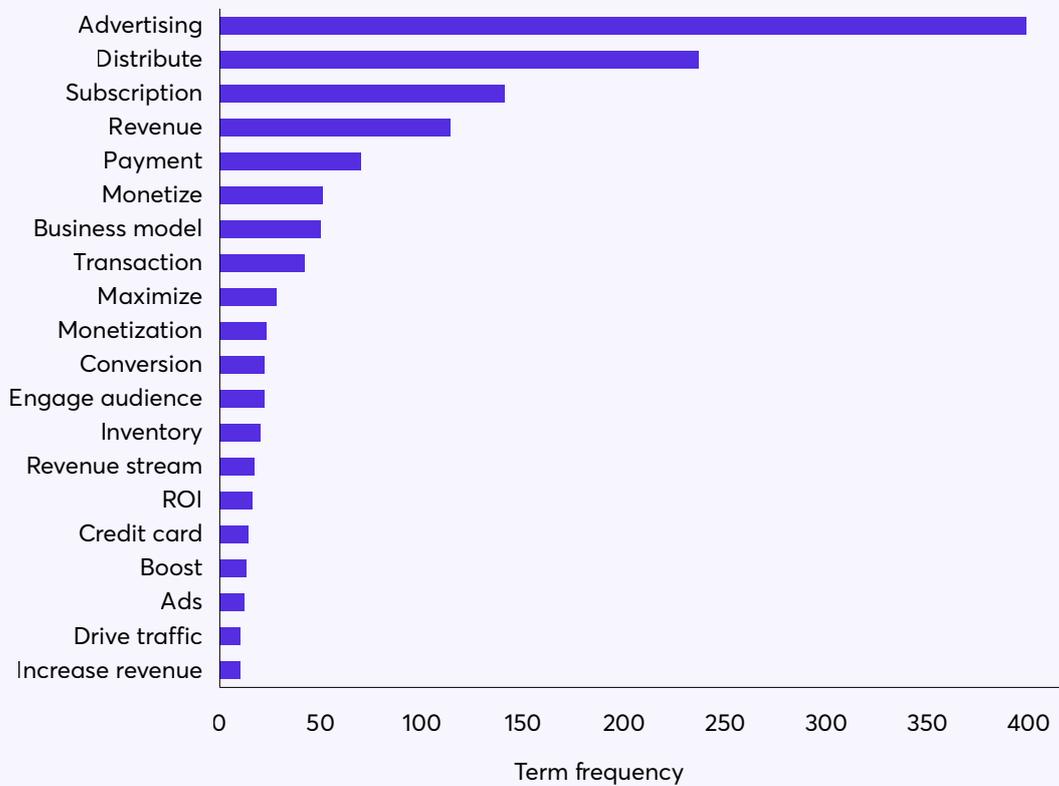
### Term frequencies for inclusive news search terms



### News revenue

ad\_exchange, ad\_inventory, ad\_revenue, ad\_server, additional\_revenue, ads, adsense, advertiser\_publisher, advertising, audience\_engagement, bidding, billing, business\_model, conversion\_rate, cryptopayment, distribute\_advertisement, distribute\_monetize, distribution\_monetization, drive\_traffic, earn\_revenue, fee, finance\_sustainability, financial\_sustainability, freemium, increase\_revenue, increase\_sale, incremental\_revenue, inventory, maximize\_revenue, micropayment, model, monetise, monetization, monetize, payment, paypal, reduce\_cost, revenue, revenue\_stream, rtb, self\_serve, subscription, subscription\_base, subscription\_fee, sustain\_financial, transaction, yield

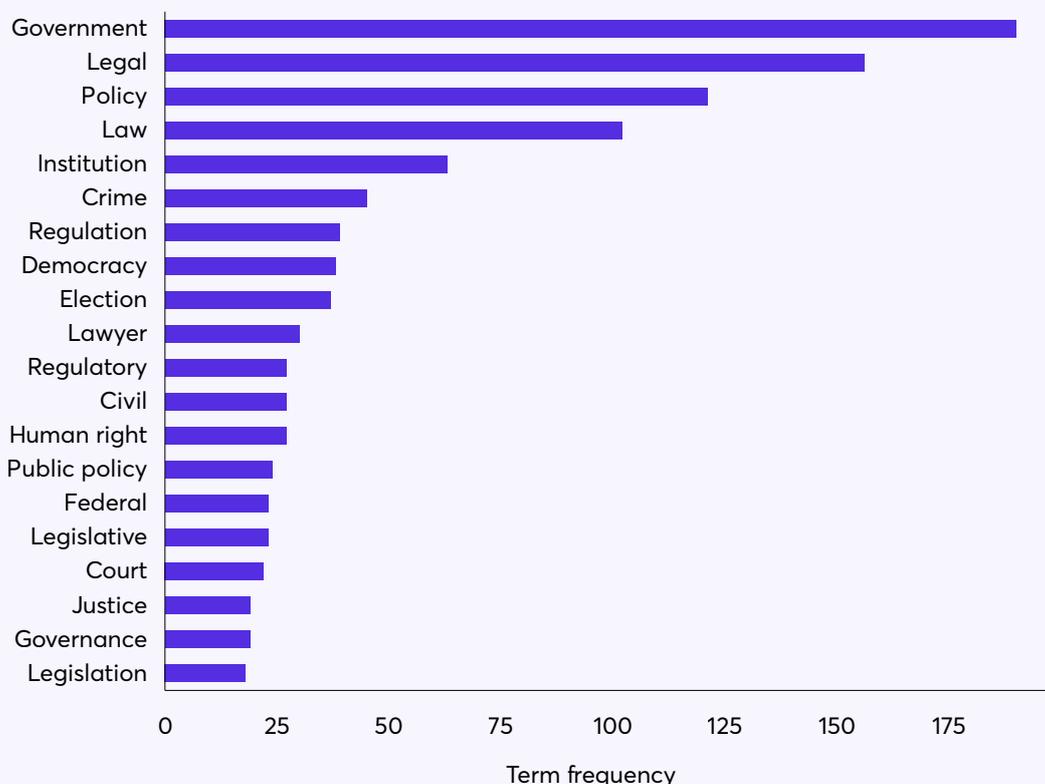
### Term frequencies for news revenue search terms



### Legal and civic news

administration, attorney, cancer, censorship, city\_county, civil, coroner, corruption, council, court, crime, criminal, democracy, election, federal, federal\_government, federal\_state, govern, governance, government, government\_agency, government\_department, government\_official, governmental, harassment, human\_right, institution, investigation, judicial, justice, justice\_system, labor, law, law\_enforcement, law\_firm, lawyer, legal, legal\_regulatory, legislation, legislative, liberty, litigation, local\_government, police, policy, prison, public\_administration, public\_policy, public\_sector, regulate, regulation, regulatory, state\_federal, state\_government, victim

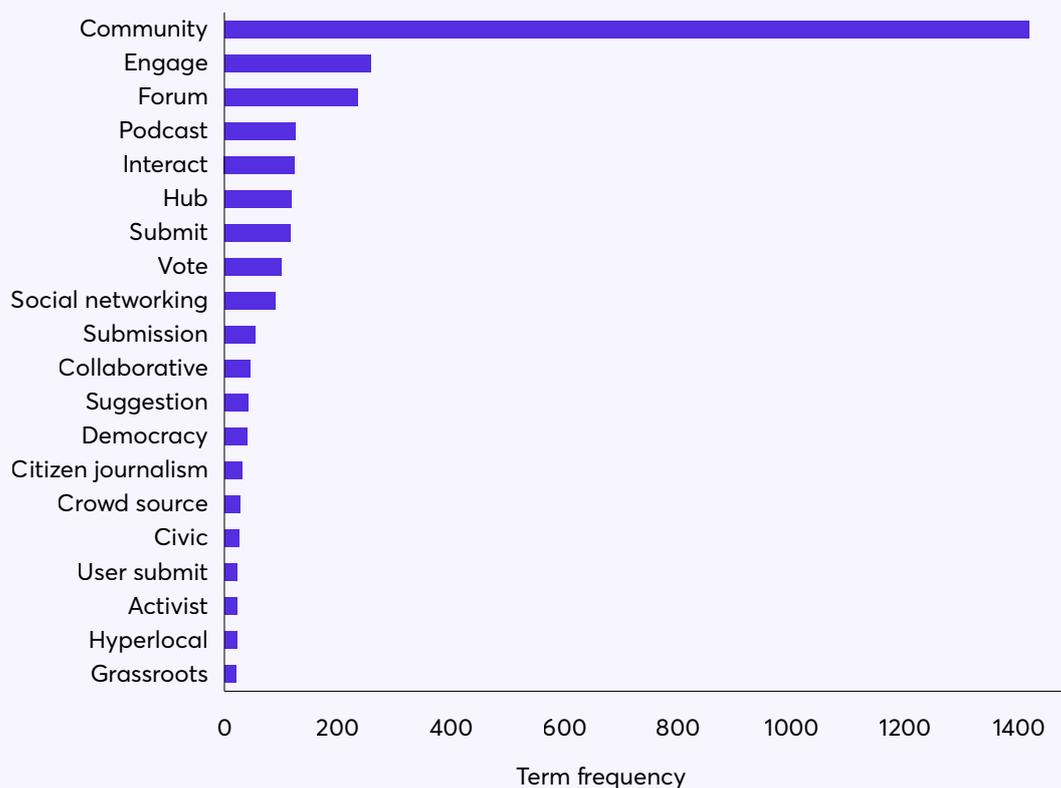
### Term frequencies for legal and civic news search terms



### News engagement

actively\_engage, activist, annotation, article\_submission, audience\_engagement, blog\_forum, bookmarking, bridge\_gap, build\_community, carbon, citizen\_journalism, citizen\_journalist, civic, civic\_engagement, co\_create, cocomment, collaborative, community, community\_newspaper, community\_people, community\_site, cooperative, cooperative\_own, create\_community, crowd\_source, crowdsourcing, deep\_engagement, discussion\_forum, engage, engage\_community, exchange\_idea, forum, grassroots, guest\_post, healthy\_lifestyle, hub, human\_right, hyperlocal, inequality, interact, interaction, journalists, membership\_organization, mentor, mentorship, message\_board, micro\_blogge, microblogging, misinformation, outreach, participatory, podcast, podcasting, post\_comment, question\_answer, reform, social\_bookmarking, social\_networking, submission, submit, submit\_article, user\_submit, voice\_hear, vote, vote\_comment, want\_hear, wiki, wikis

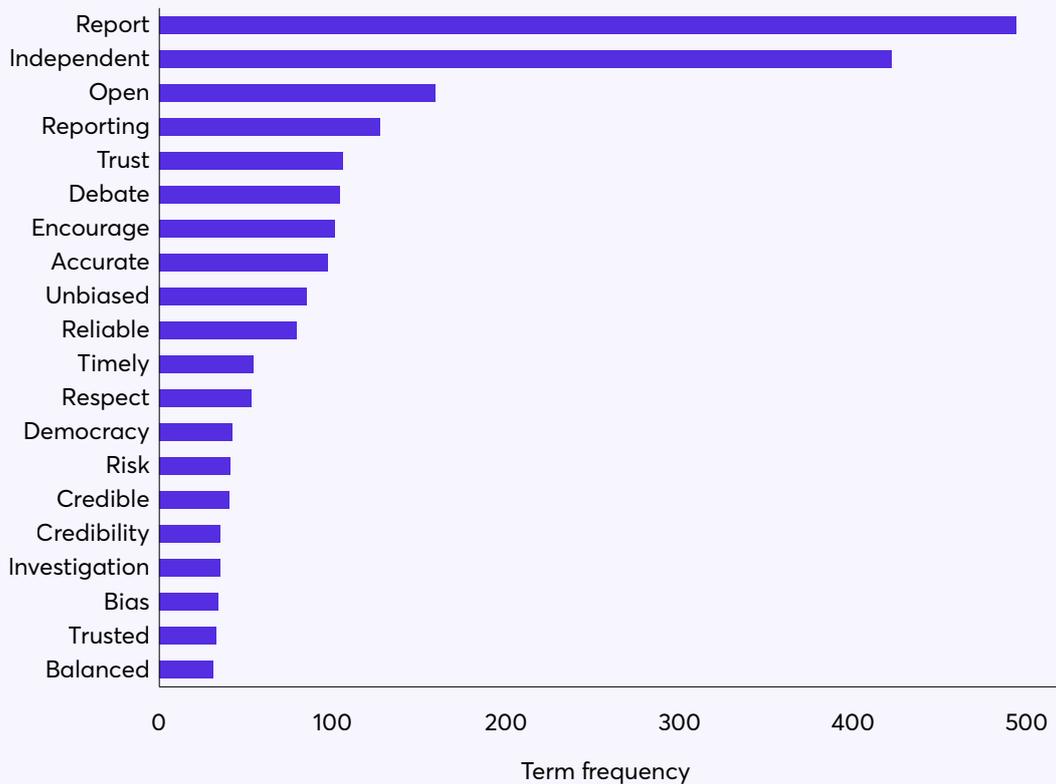
### Term frequencies for news engagement search terms



### High-quality news

accuracy, accurate, accurately, adhere, authoritative\_voice, balanced, bias, biased, bring\_transparency, civic, civil, clarity, concise, constructive, credibility, credible, credible\_source, critical, delve, democracy, democratic, depth\_analysis, depth\_reporting, developer\_designer, dialogue, dignity, disagree, discourse, disinformation, editorial\_excellence, ensure, ethic, ethical, experienced\_journalist, expertunbiased, fact\_check, factual, fair, fairness, fake\_news, hard\_hit, high\_standard, honesty, impartial, incisive, independent, integrity, investigation, investigative, investigative\_journalism, investigative\_report, investigative\_reporting, journalist\_analyst, journalistic\_integrity, misinformation, misuse, moral, non\_biased, non\_partisan, nonpartisan, objectivity, open, openness, original\_reporting, positive\_change, positive\_impact, positively, positivity, precision, professor, proficient, project\_manager, quality\_journalism, relevance, reliability, reliable, reliable\_source, remain\_committed, report, reporting\_analysis, respect, respectful, responsibility, shine\_light, standards, timeliness, timely, timely\_accurate, transparency, transparent, trust, trust\_source, trusted, trustworthy, unbiased, verified

### Term frequencies for high-quality news search terms



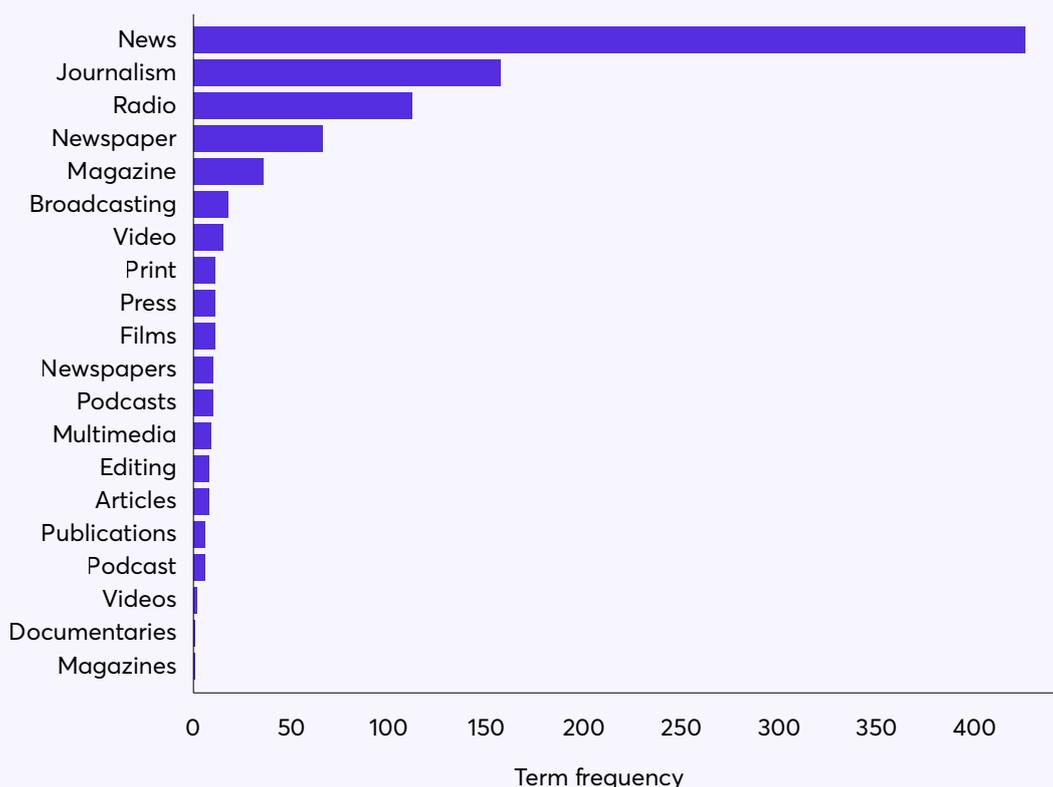
### 360Giving

This section contains lists of keywords generated through word vector keyword expansion (see Methodologies) that have been used to search for projects relating to news and journalism and benchmarking sectors (sports, health, creative arts and tech/AI) within 360Giving. Additionally, during the keyword filtering, each project description has to include at least one overarching topic in its industry. For example, a project must have "AI" or "technology" in its description.

#### News and journalism

news, journalism, podcasts, newspaper, documentaries, newspapers, podcast, articles, magazine, magazines, press, news\_stories, monthly\_newsletter, broadcast, videos, radio, quarterly\_newsletters, video, magazine, films, publications, radio, radio\_broadcasting, broadcasting, editing, print, multimedia, film\_production, graphic\_design, digital\_media

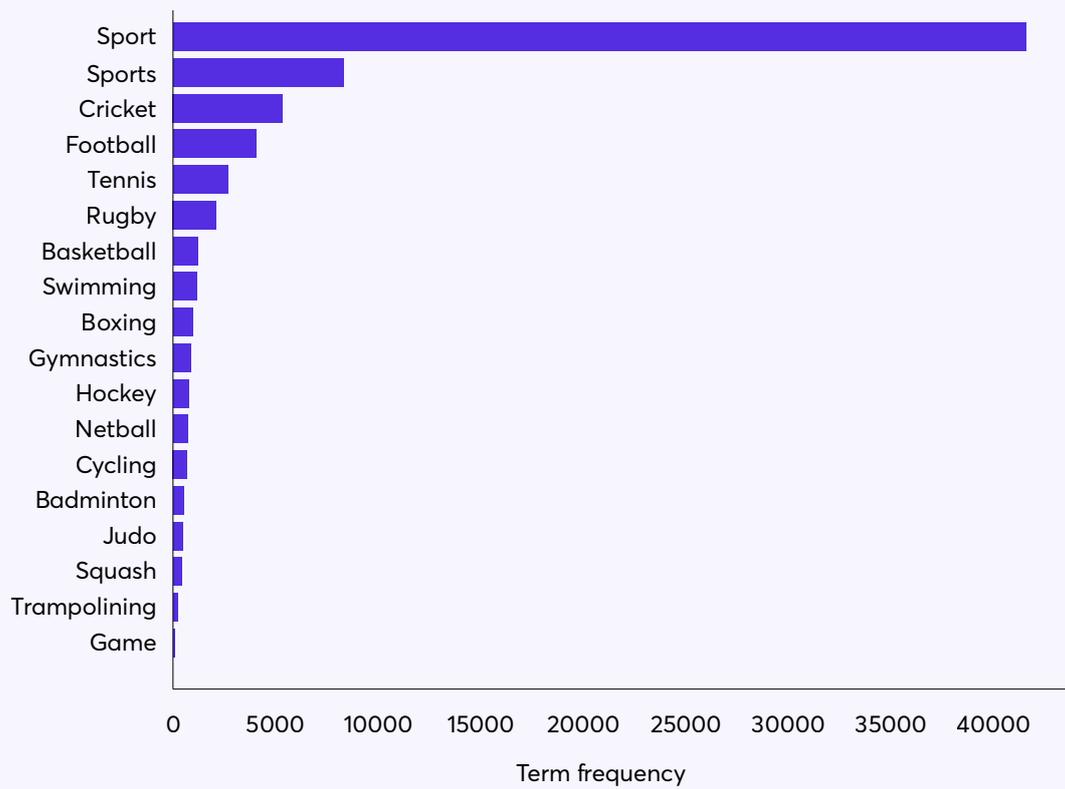
#### Term frequencies for news and journalism search terms



Sport

sports, football, sporting\_activities, tennis, hockey, cricket, basketball, boxing, rugby, cycling, swimming, gymnastics, netball, badminton, squash, judo, game, physical\_exercise, trampolining, football\_coaching

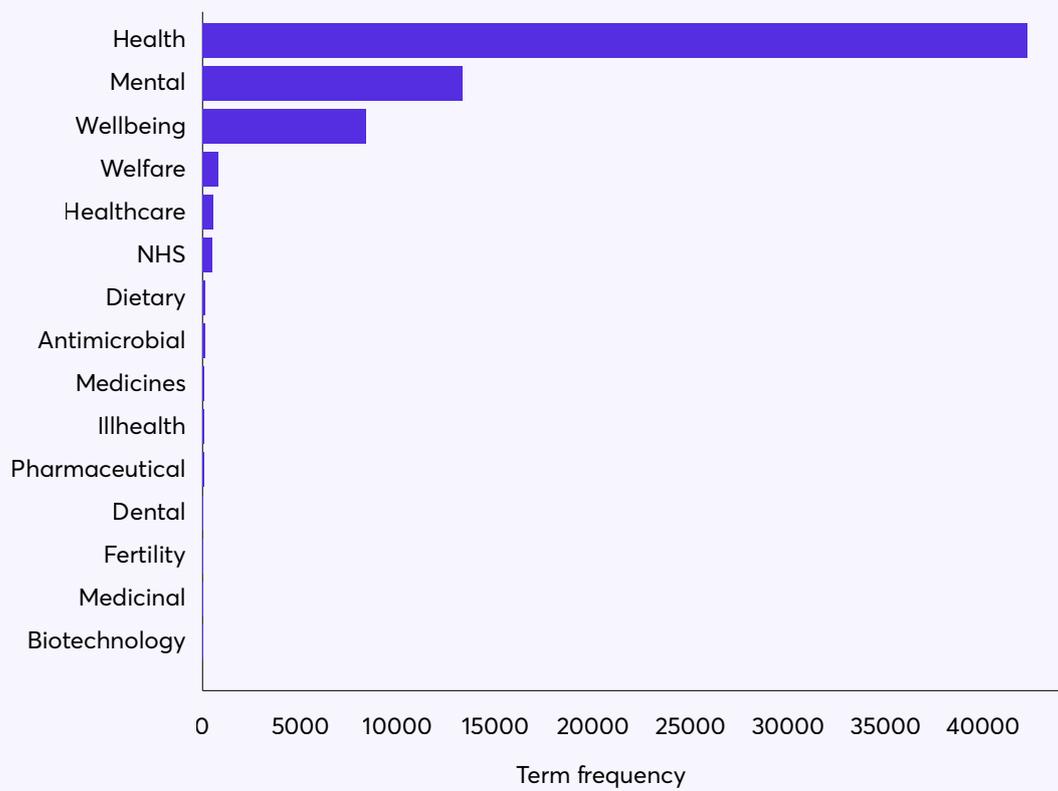
Term frequencies for sports search terms



### Health

health, health\_and\_wellbeing, health\_issues, wellbeing, mental\_health, health\_problems, ill\_health, illhealth, public\_health, psychological\_wellbeing, healthcare, mental, fertility, welfare, reproductive\_health, sexual\_health, dietary, healthy\_eating, healthy\_living, medicines, nhs, healthcare, dental, medicinal, hiv\_prevention, biotechnology, antimicrobial, pharmaceutical

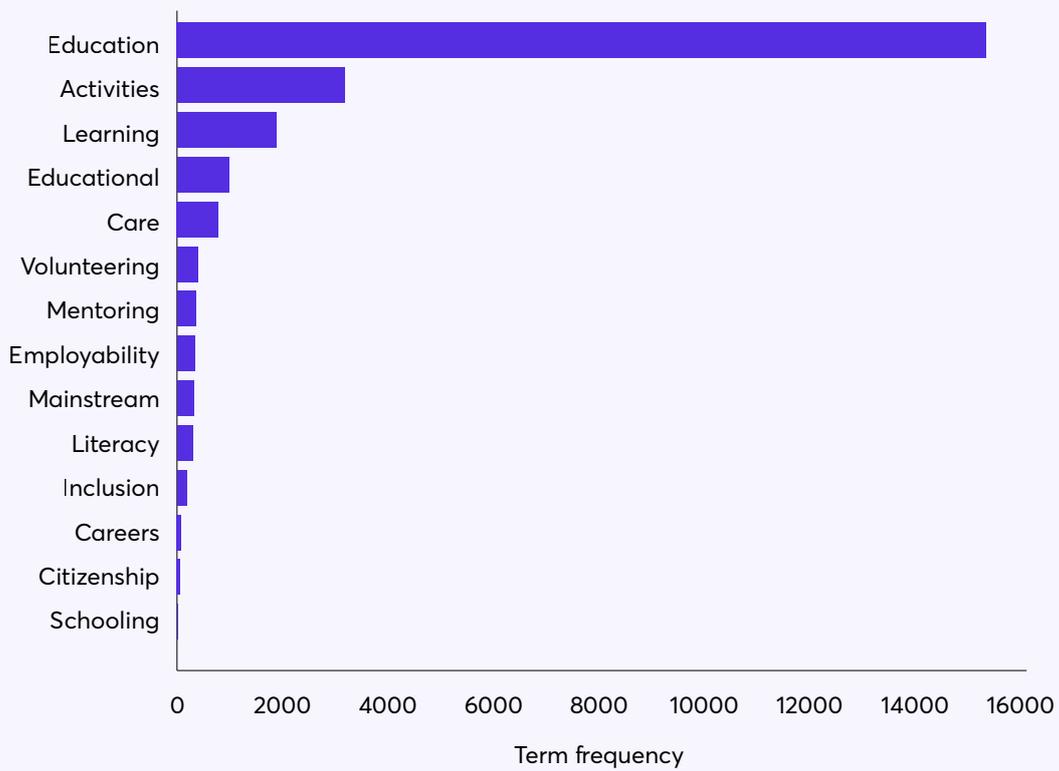
### Term frequencies for health search terms



### Education

education, educational, education\_employment, learning, higher\_education, personal\_development, mainstream\_education, citizenship, mainstream, formal\_education, schooling, care, literacy, mentoring, advance\_education, employability, inclusion, volunteering, careers, educational\_attainment, activities

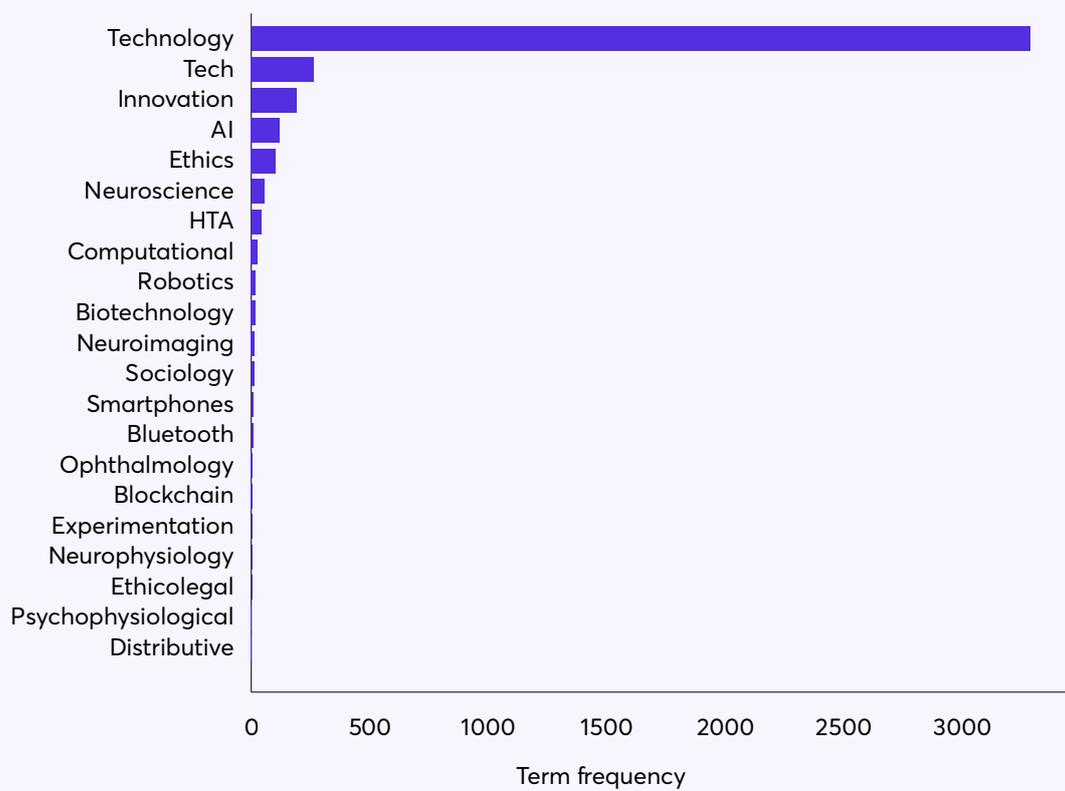
### Term frequencies for education search terms



Technology and artificial intelligence (AI)

technology, AI, psychophysiological, distributive, experimentation, ethicolegal, hta, ophthalmology, sociology, brain\_imaging, biotechnology, neuroscience, computational, adaptability, neuroimaging, blockchain, population\_genetics, mr\_imaging, gerontology, neurophysiology, smartphones, population\_genetics, ethics, technology, digital\_technology, robotics, digital\_technologies, bluetooth, smart\_phone, innovation

Term frequencies for technology and AI search terms





## Methodologies

### Word vector keyword expansion

Word vector keyword expansion is a natural language processing technique that takes a manually generated set of terms as an input and returns a larger list of related terms that exist in a set of documents. To do this a Word2Vec word embedding model (Mikolov 2013) is constructed from a set of documents (e.g. company or project descriptions). Such models are able to produce a numeric value for the approximate semantic similarity between a pair of words, based on the contexts that they have appeared in within the corpus. To perform the keyword expansion, the model is queried with each word in the manually created set of terms in turn, and the top n semantically similar terms are returned. This list is then manually filtered to remove any terms which do not fit the intended criteria. The final set of terms can then be used as a search query or filter on the original documents or a new set of documents, to obtain a list of documents relevant to the topic of interest.

The advantage of using word vector keyword expansion, rather than simply relying on a list of manually generated terms alone is that the model may surface terms that were not known or would not have been considered by an analyst. In addition, it returns only terms which are known to be present in the original corpus of documents.

### Tf-Idf document ranking

Tf-Idf (Term frequency-Inverse document frequency) is a widely used method for creating quantitative representations of unstructured text documents. The underlying algorithm transforms documents into numeric representations where the count of each word within a text is inversely weighted by the number of times it has appeared in the entire corpus. This means that words which appear in more documents are interpreted as less important than those which appear just a few times. This means that the numeric representation of each document places a greater importance on words that are more specific to that text.

Tf-Idf document ranking is a method in which Tf-Idf representations are calculated for a set of documents, taking into account only a set of keywords which are of interest. The Tf-Idf values for each document are summed and placed in rank order, with those having the highest value placed at the top. Those with higher values will contain more terms of greater importance from the original keyword list, while those at the bottom will be less relevant. The lowest ranking documents can be filtered using a manually determined threshold or other method such as a percentile threshold. In this report, we cut off all documents that have scores in the lower 10th percentile as this maximises recall, while eliminating many projects and companies which are not relevant to the search criteria.

This method is not to be confused with Tf-Idf rankings that are used to score documents based on vector similarity.

### Location quotient

We use the 'location quotient' (LQ) to capture whether activity is over-represented in a location relative to an average value (in this case usually the UK average). The higher the LQ, the more overrepresented an industry, domain or sector is. This can be used as a proxy for specialisation and local comparative advantages, and provide insight on where knowledge spillovers and innovation may be more likely to take place. Note that location quotients can get noisy when one is looking at small areas because small fluctuations in activity in one sector can have a large impact.

## Limitations

As with all studies, there are limitations to this report that should be taken into consideration when interpreting the results. The most notable limitation, which is highlighted at several points throughout the report, is the variable quality, coverage and representativeness of several of the datasets utilised. Several important funding schemes by key organisations are missing.

Some of the key omissions we are aware of include:

- Google DNI Fund, which provides €125 million over three years across the whole of Europe
- Luminate does not publish to 360Giving but is a key UK funder
- Facebook/National Council for the Training of Journalists have what appears to be a relatively untransparent initiative (no public list of all the community reporters they have funded)
- BBC Local Democracy Reporters Scheme

These data gaps limit the extent to which analyses are able to provide meaningful analyses of landscape-level trends. Understanding the reason for these gaps is an important first step in understanding how to fill them. In some cases, getting buy-in and cooperation for a common data sharing scheme amongst key actors may be an appropriate measure, however in cases of data gaps due to the fact that data are not publicly disclosed (e.g. funding data in Crunchbase), we are left with relatively little recourse. Similarly, donors may be reticent to share data.

Additionally, journalism and media are niche fields which often get subsumed into other projects (e.g. health, with the aim of getting journalists to cover a particular topic) or with Communications, making it more challenging to capture all relevant activity under a single analytic umbrella.



**nesta**

58 Victoria Embankment  
London EC4Y 0DS

+44 (0)20 7438 2500

[information@nesta.org.uk](mailto:information@nesta.org.uk)  
[facebook.com/nesta.uk](https://facebook.com/nesta.uk)  
[nesta.org.uk](http://nesta.org.uk)  
[@nesta\\_uk](https://twitter.com/nesta_uk)