



## Article

# Uncovering the State of Local News Databases in the UK: Limitations and Impacts on Research

Simona Bisiani \* and Bahareh Heravi

Surrey Institute for People-Centred AI, University of Surrey, Guildford GU2 7XH, UK; b.heravi@surrey.ac.uk

\* Correspondence: s.bisiani@surrey.ac.uk

**Abstract:** Local journalism is fundamental for a thriving democracy, yet the UK faces a decline in the number of print and digital local news outlets. Large-scale mappings of the surviving outlets offer invaluable insights to policymakers designing interventions to strengthen the sector. Due to the lack of a comprehensive national directory of UK print and digital local news outlets, researchers have resorted to datasets such as circulation auditors' databases, which have been noted to be incomplete and outdated. A lack of understanding of the magnitude of these data limitations hinders researchers from selecting optimal datasets. This study evaluates four commonly used local news databases, uncovering significant variations in their currentness and comprehensiveness. Thereafter, statistical analyses demonstrate the significant effect of each dataset's shortcomings on findings in local news research. To address this issue, triangulation and manual verification are employed to create a more comprehensive and robust dataset. This procedure generates a new national dataset of print and digital local news outlets that can be used in future research, alongside a framework for leveraging public data to build an independent research dataset. This work paves the way for more rigorous research in data-driven local news provision studies. Concluding remarks stress the importance of setting definitions and establishing clear data pipelines in an increasingly diversified and dynamic sector.

**Keywords:** print and digital local news outlets; local news provision; scale-oriented local news studies; data quality; data completeness; data currentness; spatial analysis; dataset triangulation; research dataset



**Citation:** Bisiani, Simona, and Bahareh Heravi. 2023. Uncovering the State of Local News Databases in the UK: Limitations and Impacts on Research. *Journalism and Media* 4: 1211–1231. <https://doi.org/10.3390/journalmedia4040077>

Academic Editor:  
Andreu Casero-Ripollés

Received: 2 November 2023  
Revised: 4 December 2023  
Accepted: 11 December 2023  
Published: 14 December 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

In the United Kingdom, researchers grappling with the evolution of local news ecosystems face the challenge of navigating more than four distinct datasets of print and digital local news outlets due to the absence of a national public directory of such outlets (Gulyas 2021). These datasets, each developed by different providers with unique goals, vary in size and scope. Furthermore, some of these datasets have been observed to possess inherent limitations (Gulyas 2021; PLUM Consulting 2020; Ramsay and Moore 2016). Researchers' reliance on these various datasets presents two significant concerns: first, idiosyncratic data selection compromises the comparability of research (Gulyas 2021); second, the individual limitations and intrinsic differences of these datasets are poorly understood yet crucial for ensuring precise research findings. This is particularly relevant in the context of formulating policy recommendations that tackle the insufficient provision of local news in certain regions of the country, especially during a period of substantial transformation in the local media landscape.

### 1.1. Challenges and Interventions in the UK's Local Print and Digital Sectors

The UK's local press and digital industry has seen a significant transformation in recent decades (Franklin 2006). Driven by society's shift towards digital platforms, the traditional business models of the local press are being impacted by a decline in print circulation and advertising revenue (Levy and Nielsen 2010). Despite efforts to leverage digital advertising,

local news outlets struggle against tech giants, such as Google and Facebook, who dominate this realm (Wood 2019; Picard 2003). Nationwide, these challenges have prompted measures to decrease costs and promote efficiency. These measures include the adoption of low-wage strategies, staff reductions, newsroom centralisation, and operational consolidation (Franklin 2006). The number of journalists in the local sector fell from 23,000 to 17,000 between 2007 and 2017 (Mediatique 2018). Meanwhile, the number of local newspapers in the country decreased to 1286 in 2005 from 1687 in 1985 (Franklin 2006), with industry journal Press Gazette reporting on a further 265 closures between 2005 and 2020, in terms of both print and digital (Tobitt 2020). This is particularly concerning as print and digital local news outlets in the UK are more numerous and localised than radio and television, as there are 29 local TV stations across the UK (Mediatique 2018) and 589 radio stations, combining local commercial and community radio (Ofcom 2023).

The importance of local news to society is demonstrated by research consistently showing that local news enhances democratic engagement and community awareness (Napoli et al. 2019; The Media Reform Coalition 2021). However, the concentration of news outlets in urban areas leads to urban-centric news, neglecting rural and deprived regions and exacerbating information scarcity (Abernathy 2020; Gulyas 2021; Sjøvaag 2015). The closure of local newspapers results in diminished government oversight and increased costs (Gao et al. 2020). Areas with fewer local news outlets have reduced political competitiveness and lower voter turnout (Hayes and Lawless 2018; Rubado and Jennings 2020). In short, the absence of service journalism and unequal news distribution intensify inequalities, shape democratic processes, and impact community representation.

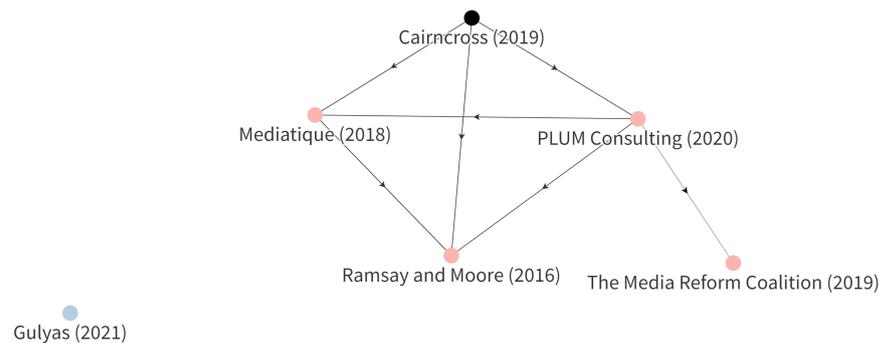
In response to the challenging dynamics facing the local news industry, the UK has commissioned research into the sustainability of the sector, which culminated in The Cairncross Review. The Cairncross Review proposes strategies to safeguard journalism's future, such as tax relief, local public-interest news funding, and further investigation into the online advertising market (Cairncross 2019; PLUM Consulting 2020). Although television and radio are two important suppliers of local news in the UK, the print and digital news sectors produce the largest amount of original local news stories, they also struggle the most in terms of generating revenue (Cairncross 2019). The UK responded to the Review in May 2023 by acknowledging the sector's strained state and the need for interventions to support local news providers, including charitable status for public interest news providers, a Digital Markets, Competition and Consumers Bill, and initiatives for the BBC to support rather than compete with local news providers (UK Parliament 2023).

Interventions to support local journalism rely on local news provision research, focusing on either scale or depth. Scale-oriented studies analyse national patterns, including the number and distribution of local news outlets, ownership diversity, and news deserts, which are geographical communities devoid of local news (Abernathy 2016; Gulyas and Baines 2020; Nygren and Schjærff Engelbrecht 2018), often using spatial visualisations (Gulyas 2021). These findings help pinpoint regions lacking resources and inform appropriate interventions. These studies address a lack of systematic evidence about the industry's success or decline, without which policymakers cannot act meaningfully (Ramsay and Moore 2016). However, scale-oriented local news research in the UK faces substantial challenges due to the lack of quality public data on local news outlets (Gulyas 2021).

### *1.2. Tracking Local News Outlets in the UK: Research Overview and Data Sources*

This section examines scale-oriented studies on local print and digital news in the UK. A systematic literature search was conducted to identify relevant studies. The search involved Scopus and the following query: ALL ("mapping local news" OR "measuring local news" OR "measuring local media" OR "mapping local media" OR "number of news outlets" OR "number of local news outlets" OR "local news deserts" AND "in the UK"). This search returned thirteen studies; one, by Gulyas (2021), was relevant, while twelve were excluded (see Appendix A for search results).

To explore research beyond academia, a citation chain search based on The Cairncross Review (Cairncross 2019) was performed. This public sector research publication was chosen for its recentness, comprehensiveness, and relevance. It led to the discovery of four additional studies. Figure 1 illustrates the citation network of these five scale-oriented studies. The relatively small network size, particularly the limited number of academic studies, reflects the emerging nature of local news provision research (Gulyas and Baines 2020). Furthermore, it suggests that conducting national-scale mappings of local news provision may present hurdles, which could be due to the lack of comprehensive, high-quality data (Gulyas 2021; Ramsay and Moore 2016; The Media Reform Coalition 2015).



**Figure 1.** The citation network structure of studies tracking the number of local news outlets built from Cairncross (2019) in red (Mediatique 2018; PLUM Consulting 2020; Ramsay and Moore 2016; The Media Reform Coalition 2019) and from the database search of Scopus in blue (Gulyas 2021).

Out of the retrieved studies, two are academic: Ramsay and Moore (2016) examined democratic deficiencies in local news provision in terms of print and digital and found a lack of coverage in many local authority districts. Gulyas (2021) focused on spatial inequalities in English local news print outlets, showcasing that over half of communities in the lowest decile for deprivation had between zero and one newspaper. The remaining studies stem from campaign groups and government-funded research associated with The Cairncross Review. The Media Reform Coalition's (MRC's) (The Media Reform Coalition 2019) work on media plurality, which was repeated in 2023, demonstrated the temporal increase in the concentration of media ownership, with six publishers owning nearly two-thirds of all local news print and digital outlets. Mediatique (2018) provided insights into structural changes in the media sector, including job losses and revenue decline in both print and digital. PLUM Consulting (2020) explored the relationship between print local news and democratic engagement.

While these studies offer valuable insights into the UK's local print and digital sector, comparing their findings is challenging due to the use of distinct datasets. Data selection tends to differ as, as Gulyas (2021) noted, there is limited availability of publicly accessible, comprehensive data on local news provision in the UK. In contrast, a few countries have implemented publicly funded initiatives to systematically monitor the dynamics of their local news industries. For instance, in the US, "The Expanding News Desert" initiative stands out for its regular monitoring of digital and print outlet closures and relocations (Abernathy 2016, 2020). In Sweden, the Institute for Media Studies maintains a database of local news outlets, which offers a reliable tool for research looking into the effects of policies on local journalism (Andersson 2020; Bisiani 2021; Institutet för Mediestudier 2023). In Australia, the News Index by the Public Interest Journalism Initiative provides interactive maps as well as monthly and quarterly reports, offering a tool to identify communities lacking public interest journalism (Dickson 2022; Public Interest Journalism Initiative 2023). These initiatives' datasets serve as indispensable tools for the continuous monitoring of the health and dynamics of the local news industry. However, the UK is not alone in lacking a local news directory. In Canada, for example, researchers are calling for an annual census of local news outlets (Lindgren 2023). In light of such a gap, the Public Interest News Foundation

(PINF), a UK-based charity with the aim to study and promote independent news, released its own dataset in June 2023 of local news outlets in the UK, encompassing print, digital, radio, and television (PINF 2023).

Due to the lack of a nationally maintained directory of print and digital local news outlets, researchers in the UK have resorted to disparate datasets that somewhat approximate and capture the publications in existence in the country, such as circulation figures from auditors such as the Joint Industry Currency for Regional Media Research (JICREG) and the Audit Bureau of Circulations (ABC). Ramsay and Moore (2016) constitute an exception, as they integrated multiple databases and reports (Table 1). The MRC, which traditionally used an independent dataset based on Ramsay and Moore’s work, now integrates the Public Interest News Foundation’s (PINF) novel local news dataset.

**Table 1.** Overview of key studies measuring local news provision and the datasets used to identify print and digital local news outlets.

Research Group	Primary Datasets Used	Dataset Status	Outlets Included
Ramsay and Moore (2016)	Local Media Works MediaTel Connected British Newspapers Online	Available Proprietary Unavailable	Those that are recorded by JICREG Unknown Unknown
Mediatique (2018)	ABC Ramsay and Moore (2016)	Proprietary Unavailable	Print and digital outlets with membership Print and digital outlets broadly
Gulyas (2021)	JICREG	Available	Print outlets with audited circulation
PLUM Consulting (2020)	JICREG	Proprietary	Print outlets with audited circulation
The Media Reform Coalition (2023)	PINF The Media Reform Coalition (2021)	Available Available	Print, digital, radio, and television outlets that are incorporated Print and digital outlets broadly

Using these datasets presents challenges, including issues of reproducibility due to data being unavailable or reliance on proprietary data. In the case of circulation auditors JICREG and ABC, data can be either retrieved freely or be purchased. The free tier contains no detailed geographical information about an outlet’s circulation, making it inutile for any spatial analysis unless the researcher manually or externally obtains geographical information about the outlet.

### 1.3. Limitations of Local News Outlet Datasets

One of the most prominent concerns arising from the absence of a dedicated research dataset of print and digital local news outlets in the UK is the presence of limitations in the datasets used to approximate the local news landscape, as highlighted in the methodological sections of the reviewed studies (Gulyas 2021; PLUM Consulting 2020; Ramsay and Moore 2016; The Media Reform Coalition 2015). Two main limitations were noticed: the presence of outdated information in the datasets, and the exclusion, both systematic and not, of news outlets. For simplicity, the extent to which datasets are up-to-date will be hereon referred to as “currentness”, while the extent to which they are comprehensive will be referred to as “completeness”.

Currentness describes the degree to which the data reflect the landscape at the point of adoption by the researcher and the degree to which the data provider maintains its current data. This is particularly relevant due to the rapid pace of newspaper closures and launches that occur in the sector, as well as extensive ownership acquisitions in the sector. This challenge was directly witnessed in the first The Media Reform Coalition (2015, p. 20) report:

“...the rate of consolidation in the UK local newspaper industry is such that one merger took place during the course of the research (Newsquest’s purchase of Romanes Newspapers in May 2015), while another—far more serious, from a plurality perspective—involving the potential purchase of the remaining share of Local World by Trinity Mirror may currently be underway at the time of writing (September 2015)”.

Completeness describes the degree to which the data cover the overall landscape of local news outlets in the UK. In this regard, datasets have been noted to be incomplete for two reasons: a quality fault in the data and the systematic exclusion of certain outlets from these datasets. Relating to the former, Ramsay and Moore (2016, p. 18) noted:

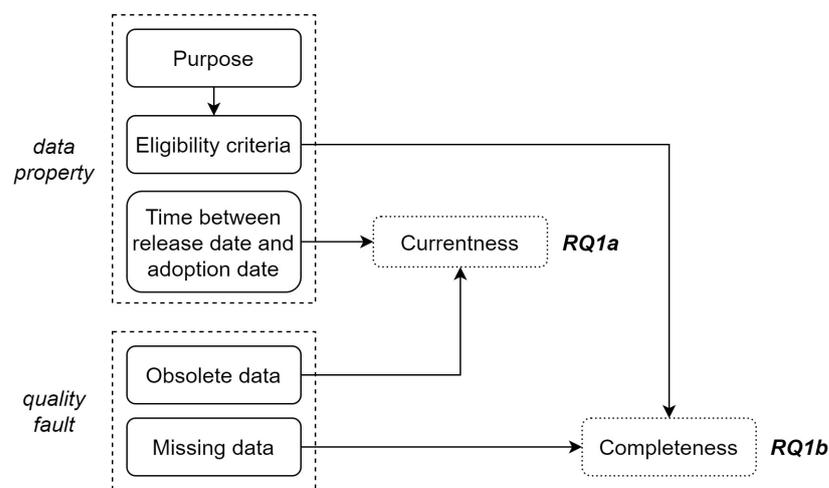
“These databases miss out a number of newspapers, and in many cases contain only grouped or aggregate information for a number of titles covering different local areas.”

Similar issues were noted by PLUM Consulting (2020, pp. 31–32), which brought the example of two papers from Ealing that were seemingly covered by JICREG but missing from the dataset:

“[...] the JICREG data for 2019 did not include either of these titles for Ealing [...]. Analysis of circulation and readership data for these titles in Mediatel contains JICREG readership data for the Ealing Gazette dated May 2019, further confusing the issue.”

Additionally, datasets tend to systematically exclude subgroups of news outlets due to the presence of eligibility criteria that determine which local news outlets satisfy the conditions to be featured in the dataset. In this regard, Gulyas (2021) noted that JICREG data could not capture outlets without circulation audits, while PLUM Consulting (2020) highlighted the absence of online-only and hyperlocal outlets.

From a conceptual perspective, these datasets’ limitations are tied to five elements: data purpose, eligibility criteria, missing data (whether the dataset contains all observations it is set to include based on the stated eligibility criteria), obsolete data (whether it contains any outdated information at the date of release), and the time gap between dataset release and adoption by the researcher. These elements interact to affect currentness and impact currentness (Figure 2).



**Figure 2.** A model of currentness and completeness.

While researchers have acknowledged issues with currentness and completeness, as the paragraphs above demonstrate, the insights and observations presented tend to be isolated, lack quantification, and apply ad hoc to the dataset used by a given researcher. There is thus a knowledge gap in terms of understanding the extent to which commonly employed local news datasets in local news provision studies are current and complete. Such evaluation would quantify the extent of the constraints and quality faults of these datasets, guiding researchers in comparing available datasets and adopting one in consideration of their intrinsic limitations.

The first objective of this research is thus to measure the limitations of the local news datasets used in the studies reviewed above by addressing the following research question: RQ1. What are the shortcomings of outlets’ datasets used in local news provision research? Two sub-questions are formulated: RQ1a. What is the degree of currentness of the dataset? RQ1b. What is the degree of completeness of the dataset?

The relevance of understanding local news datasets’ limitations is to ultimately promote a deeper grasp of the effect of the reported shortcomings on research findings. The

assumption here is that using outdated or incomplete data may lead to erroneous conclusions and unreliable policy recommendations, especially when addressing issues like news deserts and spatial inequality in terms of the supply of print and digital local news. As the UK government realises that further interventions to sustain local journalism are necessary (UK Parliament 2023), there is an urgent need to assess if and how these datasets, given their limitations, affect research outcomes. Failing to quantify the errors introduced by the dataset could lead to incorrect conclusions, particularly when most UK districts are only supplied with news by one publication or none (Gulyas 2021; Ramsay and Moore 2016). For instance, in interventions aimed at reducing news deserts, this could result in the misallocation of funds to areas already covered by news outlets. In research examining the importance of local news, biased estimates may affect the establishment of a relationship between news presence and electoral turnout.

The second objective of this study is to assess the extent of variation in research outcomes based on the dataset used. To address this, a second research question is posed: RQ2. Does data selection have a significant impact on research findings?

#### *1.4. Building a Novel Dataset of Local News Outlets Leveraging Dataset Triangulation*

In light of the limited availability of high-quality local news data in the UK, Gulyas (2021) underscores the need for a concentrated effort to address this issue. While PINF's new dataset attempts to constitute a directory of local news outlets in the country, it is uncertain whether the dataset will be maintained over time. Thus, a crucial aspect that remains unaddressed is the absence of established frameworks for a local news directory of print and digital news outlets.

While building a dataset from scratch requires extensive manual searches and verification (PINF 2023), which are often unachievable for small research teams, one promising approach to expedite the data collection is dataset triangulation. Research in the United States suggests that aggregation systems can bolster the completeness and reliability of research data in local news studies (Napoli et al. 2018). This approach was undertaken by Ramsay and Moore (2016, p. 18), who, due to "notable gaps" in "publicly available data on local newspapers", aimed to create a comprehensive list of titles by using a variety of sources.

While promising, such approaches can still deter researchers due to the resource-intensive nature of manually verifying partially matching names of local news outlets across multiple datasets, a common challenge in local news databases. Additionally, in light of considerations around data quality, researchers might still want to manually verify the completeness and currentness of the data. The research team at The Media Reform Coalition (2015, p. 20), who in their earliest report used the aggregated data by Ramsay and Moore (2016), noted that "a significant amount of research was necessary to arrive at the figures produced in this report".

This study's third objective is to test the viability and added benefit of building an independent research dataset by using data triangulation as a starting point and enhancing the dataset through manual verification and independent research. To evaluate the benefit of creating an independent dataset using such an approach, the following question is posed: RQ3. How does the novel dataset compare to the local news datasets that were used to build it? In particular, these dimensions were assessed: RQ3a. How does the completeness of the novel dataset compare to the datasets, as indicated by the number of entries missing in each dataset? RQ3b. What is the extent of unique information that each source contributes to the aggregated dataset?

To summarise, the aim of this study is to promote and facilitate sound research in scale-oriented local news provision studies with a focus on print and digital news outlets. To achieve this, the study has defined three objectives.

Objective 1: the quantification of shortcomings within local news datasets of print and digital used by local news provision researchers. This analysis will focus on measuring the degree of currentness and completeness of these datasets.

Objective 2: the determination of whether dataset selection, in consideration of datasets' shortcomings, has a significant impact on research findings.

Objective 3: the testing and evaluation of dataset triangulation and manual verification as a framework for building an independent research dataset of print and digital local news outlets in the UK.

## 2. Materials and Methods

### 2.1. Data Collection

All open datasets from the studies reviewed (Table 1) were retrieved in their latest available edition, excluding Local Media Works, as the database sources its data from JICREG, thus effectively being a copy of the latter. The datasets under investigation in this study are as follows: the UK's ABC latest circulation set for local print and digital titles, JICREG's latest free readership report, the MRC report data from 2023, and PINF's new dataset of local news outlets filtered to retain only print and digital publications. The datasets were downloaded in October 2023 (see Appendix B for location, retrieval date, and metadata), and a total of 3081 observations were accumulated.

To ensure comparability, a standard for what constitutes a print or digital local news outlet was set. Following a place-based definition (Gulyas and Baines 2020), an outlet was considered local if its primary audience was a geographical community of regional or sub-regional size, excluding nations and London (Ramsay and Moore 2016). Legacy and community print or digital media with a primary focus on news were included, recognizing the important role of hyperlocals in the local news ecosystem (Leckner et al. 2019). Newspapers published several days apart and recorded separately were merged into a single record (e.g., Lynn Tue and Lynn Fri became Lynn), while distinct sub-local editions, where identifiable, were kept separate in the recognition that they might cover different geographies (The Media Reform Coalition 2015). These decisions aim to reflect the local news ecosystem in a broad and inclusive manner while building upon previous studies.

Based upon the above, a total of twenty-one outlets that covered the whole of Scotland, Wales, and London were removed from the datasets, following Ramsay and Moore's (2016) understanding of local as more narrow than mere sub-national. Finally, observations marked as "closed" were found and removed from the PINF and MRC datasets.

### 2.2. Framework for Building a Novel Research Dataset

Given the lack of a national directory of print and digital local news outlets, the field might benefit from a framework for building an independent, high-quality research dataset. Particularly, some of the datasets under study, alongside the novel dataset built for this study, are merely snapshots of the industry at one point in time. As such, a framework would be a useful addition to the current toolkit of local news researchers. The procedure outlined here describes how the novel print and digital local news dataset was built. Such a procedure could be adopted by future researchers wishing to improve the quality of the data used in their research. The procedure consists of two stages: data triangulation and data verification.

The aim of data triangulation was to offer a convenient starting point in the process of building a novel local news outlets' dataset reflecting the current landscape of print and digital local news outlets in the UK. An intensive and time-consuming process of deduplication was carried out to merge matching records denoting the same entity, despite variant string formulations. This record linkage process brought the dataset size from 3081 to 1154 news outlets. A column for each dataset was created, denoting the presence or absence of each outlet in the respective dataset.

The goal of data verification was to identify any closed news outlets and to enhance the dataset by adding observations that were found to be missing from all datasets used in the triangulation. Missing observations were identified by reading trade reports from the Press Gazette and Hold the Front Page. The Independent Community News Network (ICNN) website, containing a list of independent publisher members, was also consulted to further identify hyperlocal publications. This procedure added 50 outlets to the dataset, bringing the final size to 1204 outlets. Of these, 134 outlets were found to have closed

(see Appendix C for more details about this procedure). A column was created for the novel dataset, and the 1070 outlets not deemed closed were marked as present in the novel dataset. The final dataset included one row per news outlet, a column for each dataset, and a column denoting whether the outlet had closed.

### 2.3. Research Design

The study starts by using statistical analysis to measure the degree of currentness (RQ1a) and completeness (RQ1b) of each dataset. To measure the variance in research findings of different datasets (RQ2), a statistical spatial analysis of the number of print and digital local news outlets was carried out. Lastly, a statistical analysis of the informational relationship between the novel dataset and each of the incorporated datasets was carried out, offering an understanding of the benefits of data aggregation and enhancement (RQ3a, RQ3b).

RQ1a. The first research question calls for the evaluation of the currentness of each dataset. Currentness quantifies the degree to which data are up-to-date. To calculate currentness, the percentage of titles in a dataset that are not considered “obsolete” was calculated:

$$\text{Currentness (\%)} = (1 - \text{Number Of Obsolete Titles In Dataset} / \text{Number Of Titles In Dataset}) \times 100 \quad (1)$$

Practically, the dataset column denoting whether an outlet had closed was used. In this context, “obsolete” relates to the data’s currentness when adopted by the researcher rather than the inclusion of outdated titles by the data provider. This is due to the difficulty in determining the precise closure date of news outlets, making it challenging to differentiate between inherent data quality issues and the natural decay of information over time.

RQ1b. In RQ1b, the aim was to assess the completeness of each dataset. Completeness measures how closely a dataset aligns with a novel or reference dataset inclusive of the broader local news landscape, representing the extent of overlap. Practically, it is measured as the percentage of titles in a dataset that also exist in the novel dataset:

$$\text{Completeness (\%)} = \text{Number Of Titles In Dataset That Are Also Present In The Novel Dataset} / \text{Number Of Titles In Novel Dataset} \times 100 \quad (2)$$

RQ2. To assess dataset variations regarding research findings, this study builds on previous research (Gulyas 2021; PLUM Consulting 2020; Ramsay and Moore 2016), which calculated average news outlet presence across UK spatial units. Two critical aspects needed definition: the meaning of “presence” in a spatial unit and the choice of spatial units.

In studying news presence, UK research has utilised two metrics: circulation, reflecting news consumption in a given area (Gulyas 2021; PLUM Consulting 2020; Ramsay and Moore 2016), and the area of coverage, describing an outlet’s target audience and reported area (The Media Reform Coalition 2023). Circulation data are not consistently available, making the area of coverage a practical choice. For this reason, this study followed Stoney’s (2023) method, which involved manual searches of news outlets’ “about us” sections for self-stated geographic areas or, when unavailable, derived from outlet names.

These coverage areas were then mapped to local authority districts (LADs). LADs are the lowest tier of local government in the United Kingdom. There are 296 LADs in England, 31 LADs in Scotland, 21 LADs in Wales, and 10 LADs in Northern Ireland, as of the latest available boundary release (Office for National Statistics 2023). LADs were chosen due to their prevalence and relevance in local news studies, as they are closely tied to local government and societal dimensions (PLUM Consulting 2020; Ramsay and Moore 2016). However, it is important to note that an outlet’s coverage may vary within a LAD or extend beyond district boundaries. Thus, it should be noted that LADs can sometimes be imprecise approximations of the geographically targeted area of the outlet. Each dataset was associated with one or several LADs, and statistics were generated: mean, median, and standard deviation of outlets per LAD, the deviation from the novel dataset for each LAD in each investigated dataset in terms of the number of news outlets covering the LAD, and the number of identified news deserts by each dataset.

Finally, pairwise Mann–Whitney U tests were conducted using the statistical software R version 4.3.1 (R Core Team 2023) to determine whether there were significant differences in the distribution of the number of news outlets across LADs between pairs of two datasets. The Mann–Whitney U test is a non-parametric statistical method typically used to measure the statistical difference between two groups when one or both groups do not follow a normal distribution (Mann and Whitney 1947). It was chosen as Shapiro–Wilk tests (Shapiro and Wilk 1965) showed that the distributions were significantly non-normal for all datasets, including the novel dataset ( $W = 0.85, p < 2.2 \times 10^{-16}$ ), ABC ( $W = 0.89, p = 9.605 \times 10^{-16}$ ), JICREG ( $W = 0.89, p = 2.641 \times 10^{-15}$ ), PINF ( $W = 0.88, p = 4.522 \times 10^{-16}$ ), and the MRC ( $W = 0.83, p < 2.2 \times 10^{-16}$ ). To determine which groups are statistically different from each other, the adjusted  $p$ -values derived from the Mann–Whitney U test were evaluated. Groups with adjusted  $p$ -values below a predetermined significance level of 0.05 were considered statistically different from each other. The smaller the adjusted  $p$ -value, the stronger the evidence that the observed differences were not due to random chance.

RQ3a. The informational advantage of the novel dataset over each dataset was evaluated by measuring the proportion of publications present in the novel dataset but absent in individual datasets:

$$\text{Missing Entries (\%)} = (1 - \text{Number of Titles in Dataset} / \text{Number of Titles in Novel Dataset}) \times 100 \quad (3)$$

This metric essentially flips the coin of the completeness metric, shifting the focus on the novel set advantage over each dataset. A higher value indicates that the novel dataset contains more information compared to that source.

RQ3b. The proportion of titles in the novel dataset that could only be sourced from one specific dataset was also determined, assessing the unique value of each dataset (RQ3b).

### 3. Results

#### 3.1. Currentness and Completeness

Analysis of the four local news datasets revealed a divergence in performance between currentness (RQ1a) and completeness (RQ1b) across datasets (Table 2). The MRC dataset was ranked as the most comprehensive dataset, with a completeness score of 74.13%, meaning nearly three-thirds of the outlets in the novel dataset could be found in it. However, it was the dataset containing the largest number of obsolete news outlets, with a currentness score of 89.50%, suggesting that more than one in ten of the MRC news outlets have actually closed. The PINF dataset had a currentness score of 97.71%, but its completeness score was much lower than MRC, at 63.82%. ABC and JICREG fared similarly. Both had high currentness scores (99.22% and 97.82%, respectively), but substantially lower completeness scores (33.50% and 31.80%, respectively).

**Table 2.** Results from the currentness and completeness analyses.

Dataset	Number of Outlets	Currentness (%)	Completeness (%)
Novel Dataset	1070	100	100
ABC	490	100	100
JICREG	471	99.22	99.22
MRC	1186	33.50	33.50
PINF	928	97.82	97.82

The findings provide a granular understanding of data currentness and data completeness of local news datasets. In particular, circulation auditors ABC and JICREG exhibit substantial omissions of outlets in their datasets, around two-thirds of the outlets in the novel dataset. This is an important consideration for researchers concerned with producing a real estimate of local news provision in the country. Conversely, larger datasets like MRC and PINF capture a greater share of news outlets, resulting in higher completeness scores, but they also include a higher proportion of obsolete data. This implies a trade-off between dataset size and data quality, with larger datasets potentially sacrificing currentness for completeness.

### 3.2. Variance in Research Findings

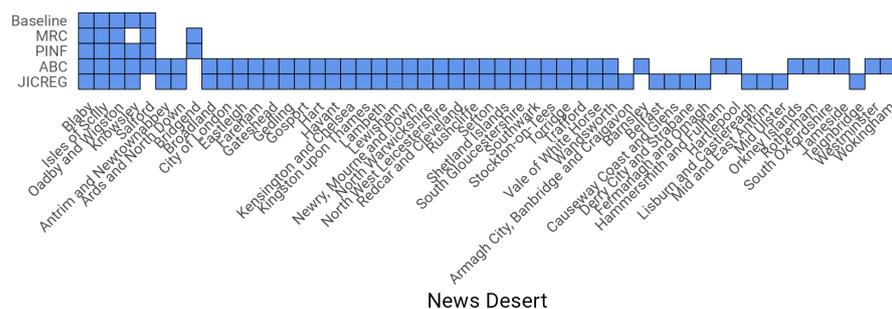
The results from the spatial analysis were used to draw conclusions about the impact of data choice on research findings. The results addressed the following questions: Which LADs are considered news deserts, according to each dataset? Compared to the novel dataset, are there any consistent patterns or over- or underrepresentation of the number of outlets in LADs? Are there significant differences across datasets in terms of the number of outlets per LAD?

Table 3 shows the mean, median, and standard deviation of the number of outlets per LAD in each dataset. The table shows that the ABC and JICREG datasets underestimate the number of outlets per LAD compared to the novel dataset. This is likely driven by the fact that these two datasets are of smaller size. The MRC and PINF datasets more closely follow the distribution of the novel dataset but there is still some variability, with PINF having a slightly lower mean and standard deviation, while the MRC reports slightly higher than the novel on these two metrics.

**Table 3.** The mean, median, and standard deviation of the number of outlets per LAD.

Dataset	Number of LADs Covered	Mean	Median	Standard Deviation
Novel Dataset	356	4.60	4	3.10
ABC	318	2.17	2	1.70
JICREG	319	2.17	2	1.67
MRC	356	5.11	4	3.61
PINF	355	4.18	4	2.76

The measurement of news deserts revealed marked differences in the number and distribution of news deserts across datasets (Figure 3). There are only three districts that all datasets agree are news deserts: the Isles of Scilly, Oadby and Wigston, and Blaby. The novel dataset only includes two further news deserts: Knowsley, and Salford, but each of these is not considered a news desert in at least one of the other datasets. PINF and MRC identify 5 and 6 news deserts, respectively, whereas JICREG and ABC count 42 and 43, respectively.

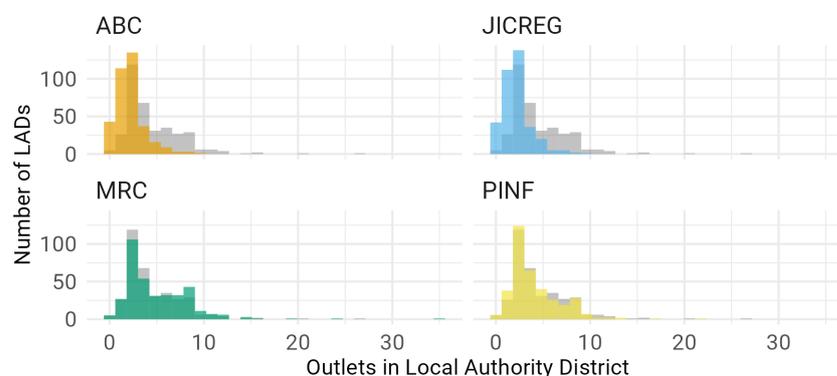


**Figure 3.** The news deserts identified by each dataset.

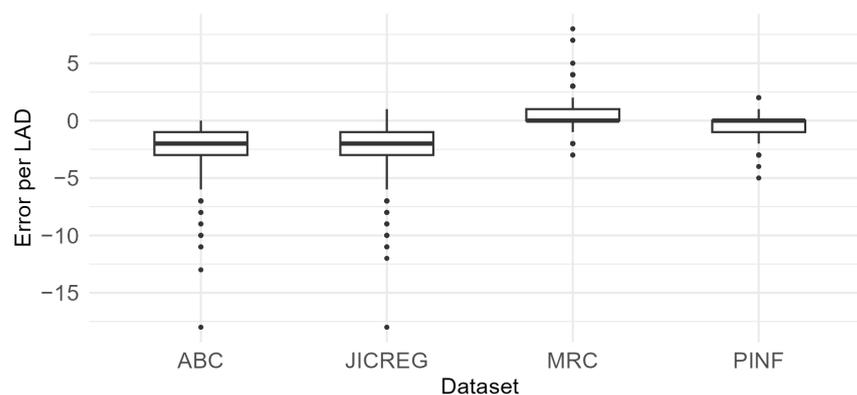
Figure 4 shows the distribution of LADs by number of outlets within each, compared against the novel dataset. The image reveals that ABC and JICREG tend to underestimate the number of outlets in LADs. On the other hand, the MRC and PINF follow more closely the distribution of the novel dataset.

This trend is more closely observed in Figure 5, which shows the error distribution by dataset, calculated as the difference between the number of outlets per LAD in a dataset and the novel dataset. The median error is displayed as the line in the middle of the boxplot, while the rectangular box shows the interquartile range. The black dots describe outliers. The picture shows that all datasets, with the exception of the MRC, undercount the number of outlets within LADs. This trend is larger for ABC and JICREG. On the other hand, PINF and the MRC have error rates that, on average, are very close to 0. In most LADs, the MRC

tends to mostly provide over-inflated estimates, albeit by a small margin. PINF, on the other hand, provides slightly under-inflated estimates.



**Figure 4.** The distribution of LADs based on the number of outlets within them in the ABC dataset (in orange), JICREG (in blue), MRC (in green), PINF (in yellow), stacked against the novel dataset (in grey).



**Figure 5.** A box plot showing the distribution of the differences in count of titles within each LAD by dataset.

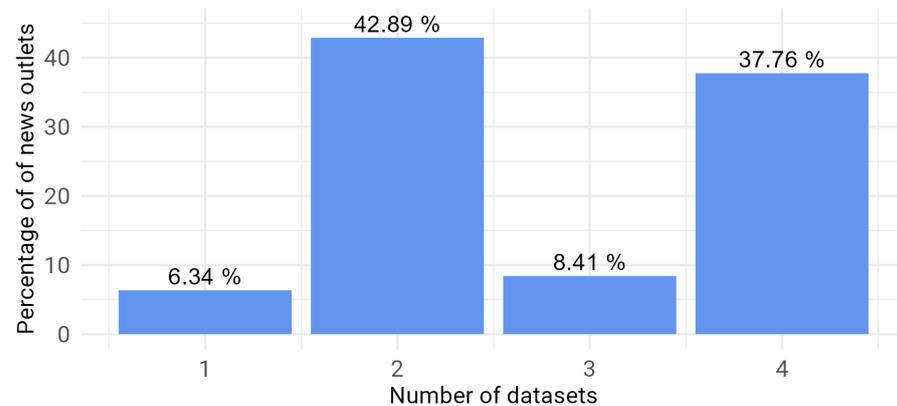
Finally, the Mann–Whitney U tests shed some light on whether these differences are statistically significant. All pairs of datasets reported statistically significant differences at a significance level of  $\alpha = 0.05$ , with three exceptions (Table 4). No statistically significant differences were observed between ABC and JICREG, between the novel dataset and MRC, and between the novel dataset and PINF.

**Table 4.** The results from the Mann–Whitney U tests. \*\*: Signifies a  $p$ -value less than 0.01, indicating a statistically significant result. \*\*\*\*: Represents a  $p$ -value less than 0.0001, indicating an extremely high level of statistical significance.

Pair	$p$ -Value Adjusted
ABC–Novel Dataset	$3.10 \times 10^{-37}$ ****
ABC–JICREG	$9.18 \times 10^{-1}$ not significant
ABC–MRC	$6.28 \times 10^{-42}$ ****
ABC–PINF	$1.40 \times 10^{-29}$ ****
Novel Dataset–JICREG	$3.13 \times 10^{-37}$ ****
Novel Dataset–MRC	$1.76 \times 10^{-1}$ not significant
Novel Dataset–PINF	$1.76 \times 10^{-1}$ not significant
JICREG–MRC	$6.28 \times 10^{-42}$ ****
JICREG–PINF	$1.48 \times 10^{-29}$ ****
MRC–PINF	$1.00 \times 10^{-3}$ **

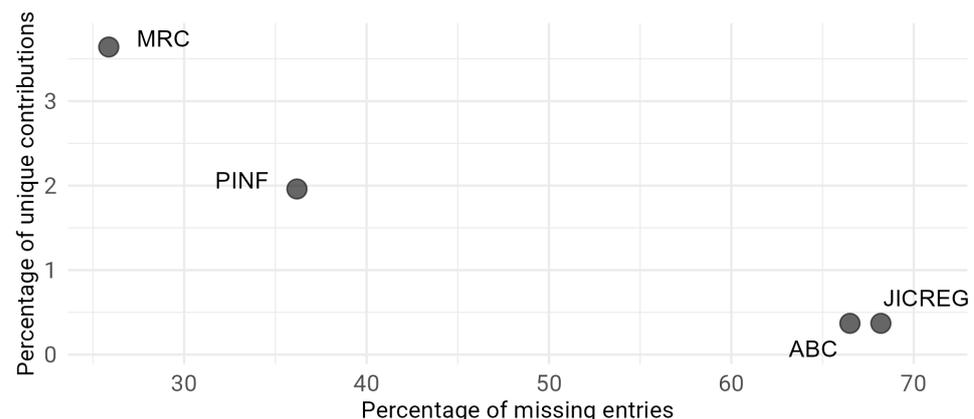
### 3.3. Dataset Triangulation

At 1070 observations, the novel dataset contains more news outlets than any other dataset, with the exception of the MRC dataset. The largest share (42.89%) of these observations can be found in two of the datasets used for the triangulation (Figure 6). Comparatively, only 37.76% of observations can be found in all four datasets.



**Figure 6.** Distribution of co-occurrences of titles in number of datasets.

Breaking down the observations that solely occur in one dataset (6.34%), it becomes evident that the highest percentages of unique contributions (RQ3b) come from the MRC and PINF. Figure 7 integrates these insights in the context of missing entries (RQ3a). The first takeaway is that MRC constitutes the most useful dataset due to a combination of a high number of unique contributions (3.64% of outlets in the novel dataset were only found in the MRC) and a low number of missing entries (25.9%). The PINF dataset is the second most informative dataset in terms of unique contributions (1.96%) but misses out on a considerable number of titles (36.2%). The limited unique amount of information provided by the ABC and JICREG datasets (0.37% each) becomes even more pronounced when considering the substantial proportion of missing entries in these datasets, amounting to 66.5% and 68.2%, respectively.



**Figure 7.** Missing entries (RQ3a) and unique contributions (RQ3b) of the four datasets.

## 4. Discussion

### 4.1. Addressing Datasets' Limitations and Divergences in Research Findings

This study set out to conduct a comprehensive analysis of the properties and limitations of four commonly employed datasets of print and digital local news outlets within scale-oriented studies of local news provision. The research was prompted by a number of reports of limitations in the datasets used (Gulyas 2021; PLUM Consulting 2020; Ramsay and Moore 2016; The Media Reform Coalition 2015) and the lack of availability of public

comprehensive quality data (Gulyas 2021). The results from the various analyses are here situated in the broader context of these datasets' characteristics and past usage in research.

Firstly, the substantial incompleteness of circulation auditor datasets may be attributed to their criteria and purposes. Both the ABC and JICREG, despite being the foundation of several local news provision studies (Gulyas 2021; Mediatique 2018; PLUM Consulting 2020), are in fact likely to exclude a large number of outlets due to their numerous inclusion criteria. The ABC dataset, requiring outlets to undergo auditing, poses the risk of excluding smaller or newer entities that may lack the resources to undergo such auditing. Similarly, JICREG's systematic exclusion of online-only outlets and hyperlocals means that the outlets missing from the dataset likely bear resemblance on various dimensions. Simultaneously, the fact that circulation auditors operate on a paid membership system could explain the higher currentness scores of ABC and JICREG, as shut-down outlets are likely naturally screened out every time ABC and JICREG re-publish their data, which occurs twice a year for the former and yearly for the latter. As the PINF and MRC datasets are built through independent research, a larger margin of error in verifying an outlet's current status can instead be expected.

In light of the observations above, should we then expect that a researcher's choice in terms of which public dataset to use will have a significant impact on their research findings? The results from the Mann–Whitney U tests prove that, more often than not, the differences observed across datasets cannot be attributed to chance alone. For researchers, this means datasets cannot be used interchangeably without anticipating an effect on research findings. Conversely, the absence of statistical significance between some datasets suggests areas of similarity or convergence. An underlying thread of similarities emerged in the analyses between circulation auditor datasets. While PINF and MRC displayed some commonalities, they still exhibited significant variations in research findings in the Mann–Whitney test. In contrast, the novel dataset showed no significant differences when compared to PINF and MRC, suggesting a closer alignment between the novel dataset and these two.

One useful application of these findings, going forward, is the provision of guidance in terms of the ability to compare studies that used datasets that were not significantly different in the Mann–Whitney test. For instance, comparing Gulyas (2021) and Mediatique (2018), both utilising JICREG and ABC data, is appropriate. In contrast, drawing comparisons between PLUM Consulting (2020) and The Media Reform Coalition (2023), which employed JICREG and the MRC dataset, may prove challenging. Even when both are built on the PINF dataset, direct comparisons of the findings in recent reports by The Media Reform Coalition (2023) and PINF (2023) may present difficulties.

On a different note, policymakers designing interventions to strengthen local media in local news-deprived areas should be aware that the identification of news deserts can vary significantly depending on the dataset used. The limited agreement across datasets on which districts across the UK are news deserts highlights the need for a more standardised dataset adoption, driven by the usage of robust data, to identify areas with inadequate news coverage. Finally, the significance of the difference between datasets suggests that policymakers should implement comprehensive assessments of methodologies from consulted research when developing interventions to address news deserts. In this light, policymakers should strive to develop targeted interventions based on accurate data and robust findings, as relying on datasets with overestimations or underestimations can result in inefficient resource allocation.

The mean and median analyses show that PINF, the MRC, and the novel dataset find that local news presence is around double what is found by ABC and JICREG. While this is significant, the disparity in the standard deviation is perhaps more interesting. Describing the spread of a distribution, the nearly double standard deviation of the novel dataset compared to the circulation auditors demonstrates the risk of underestimating local news inequality across the country when using data from circulation auditors. That is, the differential extent to which people across the UK are supplied with local news.

In terms of data aggregation, the literature has pointed towards the potential of collation systems to improve the completeness and reliability of research data (Napoli et al. 2018).

While the implementation of such systems in UK local news provision research is largely unexplored, this study finds that aggregation can benefit news outlet discovery, particularly when combining datasets with high unique contributions. By leveraging the four datasets' uniqueness, the aggregation brought the risk of missing relevant titles to a minimum, resulting in a more robust and diverse dataset. However, the degree of overlap between the four datasets is generally high. This suggests that depending on the specific overlap between pairs of datasets, some of the datasets here used might be nearly completely redundant. Looking at the unique contributions of each dataset, this seems to point to ABC and JICREG, suggesting that, if resources for carrying out triangulation are limited, leaving out the circulation auditors might negatively impact outlet discovery the least. Instead, if relying on a single dataset, this suggests that relying solely on circulation auditors' data may result in an incomplete picture of the local news landscape. Overall, if building an independent research dataset is not possible, the MRC, with its high unique contributions and relatively low missing entries, is likely to be the most informative dataset for a wide range of research questions. Nonetheless, the degree of disagreement between the novel dataset and the MRC in terms of currentness demonstrates the importance of manual verification both when adopting a single dataset and when carrying out data triangulation.

Lastly, the spatial analysis of the novel dataset confirms previous research findings (Gulyas 2021): local news provision is uneven and scarce across the country, with a quarter of LADs having access to two news outlets or less (Appendix D). These results encourage future researchers to confront broader questions about inequality, representation, and the characteristics of news provision.

#### *4.2. Addressing Challenges in Building an Independent Local News Outlet Research Dataset*

The study found that building an independent research dataset was a useful and necessary exercise that is generally recommended to researchers who wish to work with high-quality local news data. However, the construction of the dataset took several months and was ultimately highly resource-intensive. Additionally, it revealed that the feasibility of building a local news dataset is not just constrained by resources but also by the presence of several additional conceptual challenges that were not found to be addressed in the literature. These included selecting an approach to determine an outlet's closure status, defining a procedure for finding missing news outlets (such as new launches), and deciding whether outlets with certain characteristics constituted a standalone outlet. Some of these are explored in more detail in the following section, in the belief that it can be beneficial for future researchers working in this area.

Firstly, a trend was noted in the rise of online outlets incorporating several print titles. This is particularly the case with outlets from two of the largest publishers, Reach PLC and NationalWorld. For example, the website SussexWorld includes 15 different print titles from East and West Sussex, including, for example, the Worthing Herald and the Bexhill and Battle Observer. Such instances were not easy to discover, as this embedded structure was not clearly or prominently displayed on the website. For researchers interested in capturing or obtaining a sense of the unique news content generated for a particular community, this introduces some unanswered questions: To what extent are the newspapers on the same online website separate? Is there a content overlap across print outlets? What does this mean for content plurality? Furthermore, even when these instances were found, it was sometimes hard to find out whether the print edition still exists or whether the print outlet has been converted to a subsection of a bigger print–online news outlet. For example, the Brechin Advertiser was found to have ceased print publication, alongside the Montrose Review and Forfar Dispatch, and to have been incorporated into the printed Angus County Press. The masthead of the print title is displayed on the news website AngusWorld, raising questions about the meaning of incorporation but also of the promptness of these websites in keeping up-to-date with structural changes in the organisation.

Secondly, sometimes an online outlet would be found to be reported by PressGazette (Turner 2022) as a standalone outlet, but it would actually be a sub-page of a large online

website. For example, Darlington Live, by publisher Reach PLC, was found to be a subsection of TeessideLive. The same could be said about MyWirral or Warrington Live. However, with news publishers launching dedicated newsletters or social media accounts for these subsections, and with the increase in the appeal for social media accounts for audiences (Newman et al. 2023), a case could be made that these outlets might be sufficiently separated as they fundamentally attract different audiences. There needs to be a discussion in the research community about what constitutes a local news outlet in this increasingly diversified and interconnected media environment. This discourse should involve a thorough exploration of elements such as audience engagement, platform specificity, the evolving dynamics of news dissemination in the digital era, and how news consumption practices in turn shape the identity and form of local news outlets.

Thirdly, laying out a procedure to determine outlets' closures demonstrated that determining an outlet's status was, due to a lack of evidence, dependent on the researcher's priorities. For the sake of consistency, this study defined a multi-step procedure to determine this. The procedure involved searching outlets in PressGazette, looking at IPSO annual statement reports, visiting listings on publisher websites, and emailing newsrooms. The approach used in this study followed a lower-bound approach, meaning that whenever an outlet could not be determined to be open, it was deemed closed. This was performed in the belief that the numerous steps undertaken to find evidence of an outlet still operating were sufficient to conclude the outlet has likely stopped operating. The low currentness score by the MRC could thus be the result of different procedures or different priorities (e.g., reporting an upper-bound number of news outlets). This study attempts, through the choices made and the resulting work pipeline, to offer an initial framework for future researchers wishing to work with independent, high-quality data who will encounter similar challenges. However, the study also demonstrates the need for open discussions about the resources and methodological approaches available to local news provision researchers.

Finally, this procedure unveiled significant inconsistencies in record-keeping across these sources, particularly concerning publisher listings and their annual reports submitted to IPSO. While quantifying the extent of these inconsistencies was beyond the scope of this study, it "also speaks to the broader dynamics of collapse and consolidation that currently plague the industry, and the culture of the larger publishers not openly accounting for or reporting the effects of their corporate shifts", as [The Media Reform Coalition's \(2023\)](#) lead researcher Thomas Chivers put it in a written exchange with the main author of this study. Such a lack of transparency acts as a reminder that, at times, the shutting down of a newspaper or digital news outlet happens in the background, away from public knowledge. This lack of transparency has multiple implications: communities are less prepared to respond to closures, researchers must rely on outdated data for local news studies on inequality and political turnout, and legislative and executive bodies may remain unaware of the true extent of local news decline and its impact on community awareness.

The above challenges ultimately demonstrate the need for and importance of a curated and comprehensive dataset of local news outlets built upon a consensus from the research community on the grey areas identified above. This is particularly the case given the ongoing lack of a superior, regularly updated, open data source that tracks outlets on a large scale, as is available in countries like Sweden, the US, and Australia ([Abernathy 2020](#); [Institutet för Mediestudier 2023](#); [Public Interest Journalism Initiative 2023](#)). Until progress in this area is achieved, either through government or research funding to sustain such an initiative, the datasets investigated in this study, despite their individual drawbacks, continue to be invaluable assets for researchers studying local news provision in the UK, particularly for researchers lacking the resources to compile and verify spatial information and outlet status. In fact, despite being smaller in size, sources like the ABC and JICREG provide the advantage of having a regular update frequency. This regularity and the historical nature of circulators' data make them particularly appealing for temporal research.

Before concluding, certain limitations of this study must be reported. It is acknowledged that differences in definitions and inclusion terms, beyond data quality, are important

elements driving the discrepancies in results across datasets, including the novel dataset, in this analysis. This study underscores the intricate interplay of various elements influencing results in scale-oriented investigations of local news provision, spanning from data quality and dataset bias to choices in spatial units, metrics, and methodological approaches for data handling. While this study strives to disentangle and quantify some of these factors, a comprehensive understanding of why and how these differences manifest across datasets may require further in-depth analysis. Additionally, while the current data provide a reliable current estimate of the number of local print and digital news outlets in the UK, future research should explore methods for expanding and maintaining the dataset over time. Innovative approaches, such as web scraping or automated systems, hold promise for ensuring data currency and completeness.

Finally, this study is limited by the fact that it solely focuses on print and digital publications, excluding radio and television. The primary factor for this is that scholarly publications of local news provision have themselves focused on print and digital (Ramsay and Moore 2016) or print (Gulyas 2021), likely because these encounter the most significant obstacles in terms of generating revenue (Cairncross 2019). This trend is not unique to the United Kingdom. As a result, recent scholarship has noted the omission of other forms of media in local news provision research, recognising that this neglect has limited our overall understanding of news deserts (Gulyas et al. 2023; Usher 2023; Wang 2023). In this study, adding radio and television outlets would have increased the amount of time needed for the data collection, corroboration, and analysis. But it is recognized that, while local news primarily originates from newspapers (Cairncross 2019), radio and television are important elements of local news ecosystems. Without accounting for them, local news provision studies will always be telling only one part of the story, and like in the ABC and JICREG datasets, finding news deserts where there might be none simply due to a lack of accounting for these media types. For this reason, it must be acknowledged that any spatial analysis conducted with our dataset will not be suitable for assessments of the broader news ecosystem. None of the studies assessed in the literature review included reports on deficiencies in radio or television datasets. This means that we are completely unaware whether the datasets oriented towards broadcast are adequate for research, and this is something future research should investigate.

Nonetheless, this study yields valuable conclusions regarding print and digital local news data, with implications for future research. Researchers utilising the investigated datasets should be diligent in maximising data quality by considering the observed limitations of the datasets. Alternatively, data aggregation proves effective in broadening access to publication titles and bolstering the validity of findings. These findings have practical significance, offering both a framework and evidence to guide researchers and policymakers in dissecting the local news sector at scale. The insights can improve future research by guiding dataset selection, improving research comparability, and informing policymakers about the need for research support, particularly the development of comprehensive, reliable data repositories. Finally, this study contributes to the field of local news provision research by providing a reliable estimate of operative local news print and digital publications in the UK and a dataset that could be adopted for further enquiries. Determining the existence of local news outlets forms the foundation for more complex inquiries, such as identifying underserved regions or investigating factors influencing news demand, provision, and consumption. Although beyond the scope of this paper, this research lays the groundwork for comprehensive investigations into local news provision and the variables affecting news demand and consumption. Given the government's recent interest in sustaining the local news sector, such inquiries are both timely and imperative.

**Author Contributions:** Conceptualization, S.B.; methodology, S.B.; validation, S.B.; formal analysis, S.B.; investigation, S.B.; data curation, S.B.; writing—original draft preparation, S.B.; writing—review and editing, S.B.; visualization, S.B.; supervision, B.H.; project administration, S.B. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data underlying this article are stored in the University of Surrey Open Research repository. The same data can be found, alongside the script used for the data manipulation and the statistical analysis, at the following URL: <https://github.com/simonabisiani/Local-News-Datasets-Triangulation> (accessed on 10 December 2023). To cite this dataset, please use: Bisiani, Simona. 2023. *Print and Digital UK Public Local News Datasets Triangulation and Analysis 2023*. <https://doi.org/10.15126/surreydata.900924>.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Appendix A

Scopus database retrieval (6th October 2023):

1. Ayala, Jesús. 2023. "Oakland's Mayan Diaspora Overcomes Language Barriers and Finds Refuge in Radio B'alam." *Journal of Radio & Audio Media* 30 (1): 139–64. <https://doi.org/10.1080/19376529.2022.2133124>.
2. Bowd, Kathryn. 2021. "Re-Focusing on the Local: News Start-Ups, Community Engagement and Social Capital." *Australian Journalism Review*. Intellect. [https://doi.org/10.1386/ajr\\_00057\\_1](https://doi.org/10.1386/ajr_00057_1).
3. Cook, Clare Elizabeth, and Coral Milburn-Curtis. 2023. "Sharing for Sustainability: Relating Independent Community News Management Practices within Digital Platforms." *Journal of Media Business Studies* 20 (4): 320–38. <https://doi.org/10.1080/16522354.2022.2162802>.
4. Gulyas, Agnes. 2021. "Local News Deserts." In *Reappraising Local and Community News in the UK*. Routledge.
5. Jenkins, Joy, and Elia Powers. 2023. "Listening, Reflecting, and Protecting: Empathy as a Priority for Local Journalists Covering Underserved Communities." *Journalism Practice* 0 (0): 1–18. <https://doi.org/10.1080/17512786.2023.2202648>.
6. Matthews, Rachel. 2023. "Introduction: Perspectives on the Local Newspaper." In *Local Journalism*. Routledge.
7. Matthews, Rachel, and Guy Hodgson. 2023. *Local Journalism: Critical Perspectives on the Provincial Newspaper*. London: Routledge.
8. Negreira-Rey, María-Cruz, Jorge Vázquez-Herrero, and Xosé López-García. 2023. "No People, No News: News Deserts and Areas at Risk in Spain." *Media and Communication* 11 (3): 293–303. <https://doi.org/10.17645/mac.v11i3.6727>.
9. Olsen, Ragnhild Kr, and Birgit Røe Mathisen. 2023. "Deserted Local News: Exploring News Deserts From a Journalistic Recruitment Perspective." *Media and Communication* 11 (3): 390–400. <https://doi.org/10.17645/mac.v11i3.6738>.
10. Ramos, Giovanni, Luísa Torre, and Pedro Jerónimo. 2023. "No Media, No Voters? The Relationship between News Deserts and Voting Abstention." *Social Sciences* 12 (6): 345. <https://doi.org/10.3390/socsci12060345>.
11. Stonbely, Sarah. 2023. "What Makes for Robust Local News Provision? Structural Correlates of Local News Coverage for an Entire U.S. State, and Mapping Local News Using a New Method." *Journalism and Media* 4 (2): 485–505. <https://doi.org/10.3390/journalmedia4020031>.
12. Wahl-Jorgensen, Karin. 2022. "Local Knowledge and Epistemic Authority in Entrepreneurial Journalism." *Digital Journalism* 0 (0): 1–15. <https://doi.org/10.1080/21670811.2022.2128388>.
13. Warf, Barney. 2023. *Post-Truth Geographies*. Berlin: Walter de Gruyter GmbH & Co KG.

## Appendix B

**Table A1.** Supplementary information about the datasets investigated.

Data	Stated Eligibility and Inclusion Criteria	Release Date	Frequency of Updates	Date Retrieved	URL
ABC	<p>“The publication must be eligible to report under the Reporting Standards</p> <p>(a) You may apply to register a print or digital publication under the Regional Publication/ROI Newspaper Reporting Standards if it is: (i) Published in a defined region of the UK (i.e., not one that is available throughout all or most of the UK); or (ii) Published with the majority of its circulation in the Republic of Ireland. (iii) Clear from the publication name and/or content that it relates to a defined region of the UK or the Republic of Ireland” (ABC 2023, p. 4)</p>	August 2023	Twice a year	5 October 2023	<a href="https://www.abc.org.uk/data/data">https://www.abc.org.uk/data/data</a> (accessed on 5 October 2023)
JICREG	<p>“Qualifying Newspapers</p> <p>i. The JICREG Newspaper audit rules and guidelines apply to regional and local newspapers which circulate in England, Scotland, Wales, the Isle of Man, Northern Ireland and the Channel Isles and are JICREG members.</p> <p>ii. Newspapers must be published at least weekly.” (JICREG 2019, p. 5)</p> <p>“Definition of a Newspaper (also referred to as a Publication or a Title)</p> <p>i. An independently audited net circulation figure is required for each Newspaper recorded on the JICREG database</p> <p>ii. For these purposes a Newspaper is usually defined as a publication in which advertising can be routinely purchased on a stand-alone basis by national advertisers and agencies (i.e., the level at which a potential national media buyer can expect to have meaningful circulation data available).” (JICREG 2019, pp. 5–6)</p>	December 2022	Yearly	5 October 2023	<a href="http://www.jicregonline.co.uk/dashboard">http://www.jicregonline.co.uk/dashboard</a> (accessed on 5 October 2023)
PINF	<p>“To be included in the local news map, outlets must:</p> <ul style="list-style-type: none"> <li>• Be legally incorporated (i.e., a registered company, charity or society)</li> <li>• Have a principal purpose for the provision of local news; and,</li> <li>• Be recently updated.” (PINF 2023, p. 23)</li> </ul>	July 2023	Undefined	5 October 2023	dataset made accessible by PINF
MRC	Unstated	October 2023	Whenever a new report is published	5 October 2023	<a href="https://research.gold.ac.uk/id/eprint/34156/">https://research.gold.ac.uk/id/eprint/34156/</a> (accessed on 5 October 2023)

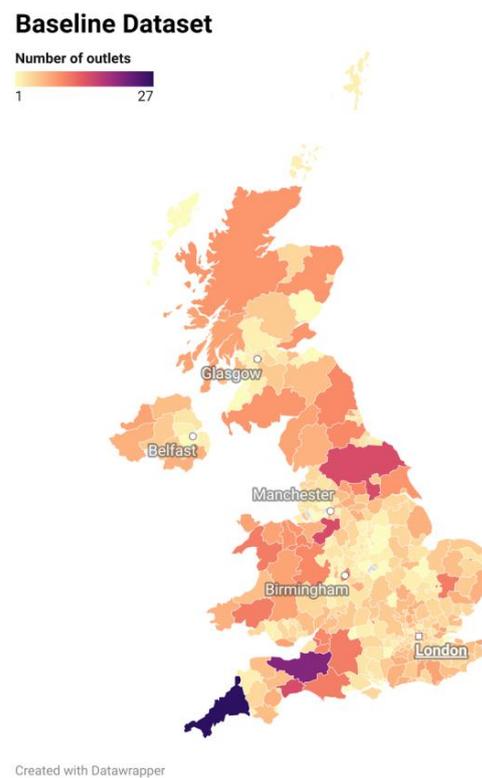
## Appendix C

Manually tagging whether the outlet is now closed was determined through these steps:

1. Look up website;
2. Look up print edition on PressReader (<https://www.pressreader.com/>) (accessed on 5 October 2023), Magzter (<https://www.magzter.com/>) (accessed on 5 October 2023), Readly (<https://gb.readly.com/>) (accessed on 5 October 2023), National World Local Subs (<https://www.localsubsplus.co.uk/>) (accessed on 5 October 2023), and Reach PLC Local Subs (<https://www.reachsubs.co.uk/>) (accessed on 5 October 2023);
3. Look up brand listings for Reach PLC (<https://www.reachsolutions.co.uk/brands>) (accessed on 5 October 2023), Newsquest (<https://www.newsquest.co.uk/news-brands>) (accessed on 5 October 2023), Tindle (<https://tindlenews.co.uk/portfolio/>) (accessed on 5 October 2023), Illiffe (<https://www.iliffemedia.co.uk/portfolio/publications/>) (accessed on 5 October 2023), Bullivant Media Ltd. (<https://bullivantmedia.com/portfolio/newspapers/>) (accessed on 5 October 2023);
4. Search through pie charts of ABC 2022/2023 certificates (these data have been extracted and stored in a dataset);

5. Search for reports on the outlet on HTFP or Press Gazette;
6. Search IPSO annual statements for the owning company of an outlet and search outlet in the report (<https://www.ipso.co.uk/monitoring/annual-statements/>) (accessed on 5 October 2023);
7. Search HTFP and Press Gazette websites for any additional reports of closures between 2022 and 2023;
8. Still no evidence of whether the outlet is still active? Then it is considered likely to have closed down.

## Appendix D



**Figure A1.** The number and distribution of local news outlets across the UK's local authority districts in the novel dataset.

## References

- ABC. 2023. ABC Regional Publications Reporting Standards. Available online: [https://www.abc.org.uk/Docs/standards/Regional\\_Publications\\_Reporting\\_Standards.pdf](https://www.abc.org.uk/Docs/standards/Regional_Publications_Reporting_Standards.pdf) (accessed on 6 October 2023).
- Abernathy, Penelope Muse. 2016. *The Rise of a New Media Baron and the Emerging Threat of News Deserts*. Chapel Hill: Center for Innovation and Sustainability in Local Media, University of North Carolina Chapel Hill: University of North Carolina School of Media and Journalism. Available online: <https://www.usnewsdeserts.com/reports/rise-new-media-baron/> (accessed on 6 October 2023).
- Abernathy, Penelope Muse. 2020. News Deserts and Ghost NEWSPAPERS: Will Local News Survive? UNC Hussman School of Journalism and Media. Available online: <https://www.usnewsdeserts.com/reports/news-deserts-and-ghost-newspapers-will-local-news-survive/> (accessed on 6 October 2023).
- Andersson, Amanda. 2020. Mediestöd för lokal journalistik: Ett Bamseplåster på ett björnbett eller en gedigen räddningsaktion i demokratins namn?: Om reaktionen på en politisk åtgärd i ett krisande medielandskap. Available online: <https://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-404002> (accessed on 6 October 2023).
- Bisiani, Simona. 2021. Blank Spots, Critical Information Needs and Local Journalism Funding. Measuring the Impact of State Financial Support for Local Journalism on the Critical Information Needs of Citizens in Swedish Journalistically Fragile Municipalities. Unpublished. Master's thesis, Linköping University, Linköping, Sweden. Available online: [https://www.researchgate.net/publication/372690433\\_Blank\\_spots\\_critical\\_information\\_needs\\_and\\_local\\_journalism\\_funding\\_Measuring\\_the\\_impact\\_of\\_state\\_financial\\_support\\_for\\_local\\_journalism\\_on\\_the\\_critical\\_information\\_needs\\_of\\_citizens\\_in\\_Swedish\\_journa](https://www.researchgate.net/publication/372690433_Blank_spots_critical_information_needs_and_local_journalism_funding_Measuring_the_impact_of_state_financial_support_for_local_journalism_on_the_critical_information_needs_of_citizens_in_Swedish_journa) (accessed on 6 October 2023).

- Cairncross, F. 2019. THE CAIRNCROSS REVIEW a Sustainable Future for Journalism. Available online: [https://assets.publishing.service.gov.uk/media/5c6bfcd4e5274a72b933311d/021919\\_DCMS\\_Cairncross\\_Review\\_.pdf](https://assets.publishing.service.gov.uk/media/5c6bfcd4e5274a72b933311d/021919_DCMS_Cairncross_Review_.pdf) (accessed on 6 October 2023).
- Dickson, Gary. 2022. Australian Newsroom Mapping Project Report: May 2022 Public Interest. Available online: <https://policycommons.net/artifacts/4397654/australian-newsroom-mapping-project-report/5194272/> (accessed on 6 October 2023).
- Franklin, Bob. 2006. *Local Journalism and Local Media: Making the Local News*. London: Routledge.
- Gao, Pengjie, Chang Lee, and Dermot Murphy. 2020. Financing dies in darkness? The impact of newspaper closures on public finance. *Journal of Financial Economics* 135: 445–67. [CrossRef]
- Gulyas, Agnes. 2021. Local News Deserts. In *Reappraising Local and Community News in the UK*. London: Routledge.
- Gulyas, Agnes, and David Baines. 2020. Introduction: Demarcating the field of local media and journalism. In *The Routledge Companion to Local Media and Journalism*. London: Routledge.
- Gulyas, Agnes, Joy Jenkins, and Annika Bergström. 2023. Places and Spaces Without News: The Contested Phenomenon of News Deserts. *Media and Communication*, 11. [CrossRef]
- Hayes, Danny, and Jennifer L. Lawless. 2018. The Decline of Local News and Its Effects: New Evidence from Longitudinal Data. *The Journal of Politics* 80: 332–36. [CrossRef]
- Institutet för Mediestudier. 2023. Kommandatabasen. Available online: <https://kommundatabas.mediestudier.se/> (accessed on 6 October 2023).
- JICREG. 2019. Minimum Audit Rules and Guidelines for Calculating and Reporting Newspaper Net Circulation Data. Available online: <https://www.jicreg.co.uk/documents/JICREGAuditRulesOctober2019-3.pdf> (accessed on 6 October 2023).
- Leckner, Sara, Carina Tenor, and Gunnar Nygren. 2019. What About the Hyperlocals?: The drivers, organization and economy of independent news media in Sweden. *Journalism Practice* 13: 68–89. [CrossRef]
- Levy, David A. L., and Rasmus Kleis Nielsen. 2010. *The Changing Business of Journalism and Its Implications for Democracy*. Oxford: University of Oxford, Reuters Institute for the Study of Journalism.
- Lindgren, April. 2023. Librarians Can Help Create Inventories for Community News Access in Canada | J-Source. November 21. Available online: <https://j-source.ca/librarians-can-help-create-inventories-for-community-news-access-in-canada/> (accessed on 6 October 2023).
- Mann, Henry B., and Donald R. Whitney. 1947. On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other. *The Annals of Mathematical Statistics* 18: 50–60. Available online: <http://www.jstor.org/stable/2236101> (accessed on 6 October 2023). [CrossRef]
- Mediatique. 2018. Overview of Recent Dynamics in the UK Press Market. Available online: [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) (accessed on 6 October 2023).
- Napoli, Philip M., Sarah Stonbely, Kathleen McCollough, and Bryce Renninger. 2019. Local Journalism and the Information Needs of Local Communities: Toward a Scalable Assessment Approach. *Journalism Practice* 13: 1024–28. [CrossRef]
- Napoli, Philip M., Matthew Weber, and Kathleen McCollough. 2018. An Approach to Assessing the Robustness of Local News Provision. In *The Routledge Handbook of Developments in Digital Journalism Studies*. London: Routledge.
- Newman, Nic, Richard Fletcher, Kirsten Eddy, Craig T Robertson, and Rasmus Kleis Nielsen. 2023. Reuters Institute Digital News Report 2023. Available online: [https://policycommons.net/artifacts/4164711/digital\\_news\\_report\\_2023/4973510/](https://policycommons.net/artifacts/4164711/digital_news_report_2023/4973510/) (accessed on 6 October 2023).
- Nygren, Gunnar, and Nønne Schjærff Engelbrecht. 2018. Vita fläckar och medieskuggor—Stora skillnader i mediernas kommunbevakning. Institutet för Mediestudier. pp. 22–53. Available online: <https://urn.kb.se/resolve?urn=urn:nbn:se:sh:diva-36970> (accessed on 6 October 2023).
- Ofcom. 2023. Media Nations Report 2023. Available online: <https://www.ofcom.org.uk/research-and-data/tv-radio-and-on-demand/media-nations-reports/media-nations-2023> (accessed on 6 October 2023).
- Office for National Statistics. 2023. Local Authority Districts (May 2023) Boundaries UK BFC. Available online: <https://geoportal.statistics.gov.uk/datasets/ons::local-authority-districts-may-2023-boundaries-uk-bfc/explore?location=51.829740,-0.216990,10.05> (accessed on 6 October 2023).
- Picard, Robert G. 2003. Cash Cows or Entrecôte: Publishing Companies and Disruptive Technologies. *Trends in Communication* 11: 127–36. [CrossRef]
- PINF. 2023. Deserts, Oases and Drylands. Available online: <https://www.publicinterestnews.org.uk/map> (accessed on 6 October 2023).
- PLUM Consulting. 2020. Research into Recent Dynamics of the Press Sector in the UK and Globally. Available online: [https://assets.publishing.service.gov.uk/media/5f7b4673e90e070dec5d9e29/Plum\\_DCMS\\_press\\_sector\\_dynamics\\_-\\_Final\\_Report\\_v4.pdf](https://assets.publishing.service.gov.uk/media/5f7b4673e90e070dec5d9e29/Plum_DCMS_press_sector_dynamics_-_Final_Report_v4.pdf) (accessed on 6 October 2023).
- Public Interest Journalism Initiative. 2023. Australian News Index. Available online: <https://newsindex.piji.com.au/outlets/> (accessed on 6 October 2023).
- R Core Team. 2023. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing. Available online: <https://www.R-project.org/> (accessed on 6 October 2023).
- Ramsay, Gordon, and Martin Moore. 2016. *Monopolising Local News: Is There an Emerging Local Democratic Deficit in the UK Due to the Decline of Local Newspapers?* London: King's College London. [CrossRef]

- Rubado, Meghan E., and Jay T. Jennings. 2020. Political Consequences of the Endangered Local Watchdog: Newspaper Decline and Mayoral Elections in the United States. *Urban Affairs Review* 56: 1327–56. [CrossRef]
- Shapiro, Samuel Sanford, and Martin B. Wilk. 1965. An Analysis of Variance Test for Normality (Complete Samples). *Biometrika* 52: 591–611. [CrossRef]
- Sjøvaag, Helle. 2015. The Emergence of Metropolitan News: Shifting Concepts of Localism in Norwegian Regional Newspapers. *Nordicom Review* 36: 17–32. [CrossRef]
- Stonbely, Sarah. 2023. What Makes for Robust Local News Provision? Structural Correlates of Local News Coverage for an Entire U.S. State, and Mapping Local News Using a New Method. *Journalism and Media* 4: 485–505. [CrossRef]
- The Media Reform Coalition. 2015. Who Owns the UK Media Report. Mapping Changes in Local News 2015–2017. More Bad News for Democracy? Available online: <https://www.mediareform.org.uk/wp-content/uploads/2015/11/Mapping-changes-in-local-news-2015-2017-interactive.pdf> (accessed on 6 October 2023).
- The Media Reform Coalition. 2019. Who Owns the UK Media? Available online: <https://www.mediareform.org.uk/wp-content/uploads/2019/03/FINALonline2.pdf> (accessed on 6 October 2023).
- The Media Reform Coalition. 2021. Who Owns the UK Media? The Media Reform Coalition. Available online: [https://www.mediareform.org.uk/wp-content/uploads/2021/03/Who-Owns-the-UK-Media\\_final2.pdf](https://www.mediareform.org.uk/wp-content/uploads/2021/03/Who-Owns-the-UK-Media_final2.pdf) (accessed on 6 October 2023).
- The Media Reform Coalition. 2023. Who Owns the UK Media-2023.pdf. Available online: <https://www.mediareform.org.uk/wp-content/uploads/2023/10/Who-Owns-the-UK-Media-2023.pdf> (accessed on 6 October 2023).
- Tobitt, Charlotte. 2020. UK Local Newspaper Closures: At Least 265 Titles Gone since 2005, But Pace of Decline Has Slowed. *Press Gazette*. August 20. Available online: <https://pressgazette.co.uk/news/uk-local-newspaper-closures-at-least-265-local-newspaper-titles-gone-since-2005-but-pace-of-decline-has-slowed/> (accessed on 6 October 2023).
- Turner, Alexandra. 2022. UK Local Newspaper Closures: Launches in Digital and Print Balance Out Decline. *Press Gazette*. July 26. Available online: <https://pressgazette.co.uk/news/uk-local-newspaper-closures-2022/> (accessed on 6 October 2023).
- UK Parliament. 2023. Sustainability of Local Journalism: CMS Committee Publishes Government response—Committees—UK Parliament. Available online: <https://committees.parliament.uk/work/6536/sustainability-of-local-journalism/news/195274/sustainability-of-local-journalism-cms-committee-publishes-government-response/> (accessed on 6 October 2023).
- Usher, Nikki. 2023. The Real Problems with the Problem of News Deserts: Toward Rooting Place, Precision, and Positionality in Scholarship on Local News and Democracy. *Political Communication* 40: 238–53. [CrossRef]
- Wang, Ryan Yang. 2023. The Geography of Newspaper Circulations: A Spatial Taxonomy of “News(Paper) Deserts” in the United States. *Media and Communication* 11: 304–17. [CrossRef]
- Wood, Sam. 2019. Online Advertising in the UK. *Plum Consulting* (blog). Available online: <https://plumconsulting.co.uk/online-advertising-in-the-uk/> (accessed on 6 October 2023).

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.