**1NIC report on historical data collection for telegraph and telephone usage and access in the UK 1840-1983.**

**Abstract**

These data provide an insight into historical public access to telegraph offices between 1854-1980 and public usage of telephones for UK (England and Wales only for access). We provide data showing telegraph network coverage and accessibility and number of telegraph stations annually 1840-1980. Telegraph accessibility is defined as a telegraph office being located within 30 minutes’ walk from any geographical network entry point. Data for public usage are provided for volume of telephone calls made and telegrams sent 1855-1980 UK wide. These are used to estimate aggregate volume of telegraph use by year measured as data bytes.

**Specifications**

**Type of data**

Table

**How data were acquired**

Data for number of internal telephone calls and telegrams sent were acquired by OCR and manual data entry extracted from information collected in UK parliamentary reports. Adobe Acrobat Pro was used, and manual checks made for accuracy.

Raw data were supplied to us by third parties giving telegraph office location. We parsed, filtered and cleaned these data from an historical post office dataset. We used Microsoft Excel and Google Open Refine for data wrangling.

We geolocated location of telegraph offices with GoogleEarth Pro. Accessibility measures were calculated with ArcGis Pro and Modaccess.

**Parameters for data collection**

Our primary concern in selecting data collection sources was reliability and high level of coverage.

Our original NIC bid included eight data series compiled from various sources. For reference, the below is copied from page 2 of our bid:

1. Timeseries data from parl. reports of postmaster general giving annual timeseries UK 1870-1980:
	* # Telegrams sent
	* # Telephone calls trunk and local
	* # Telegraphic stations (number of, each year).
	* # Telephone stations (number of, each year).
2. Timeseries data from BT archives for quasi-annual timeseries UK 1850-1870:
	* # Telegrams sent
	* # Telegraphic stations (number of, each year).
3. Timeseries data annual timeseries UK 1850-1980:
	* # Data load for telegrams converted to bytes (GB of data consumed), compatible with current OFCOM series.
4. A new decadal network coverage metric for telegraph similar based on current mobile phone coverage (OFCOM 2020 Coverage Obligations - Notice of compliance verification methodology). As this is a more complex operation, the methodology is detailed below.

We succeeded in providing all eight series. Although were unable to collect pre-1870 data (from the BT Archives) to cover b). This was in consequence of the national closure of all archives in early 2021 after the UK government response to the pandemic. Instead, we estimate data for the pre-1870 period instead using available online sources of information (telegram volume and telegraph offices numbers).

**Description of data collection**

**Primary data sources**

Early telegraph offices in 1862 were georeferenced from cartographic material obtained from Distant Writing (<https://distantwriting.co.uk/TelegraphStations1862.html>) and completed with published primary and secondary material listing telegraph offices for each company before nationalisation.

Before 1870 volume data are - for 1868 – from Wenzlhuemer, R. (2012). *Connecting the Nineteenth-Century World: The Telegraph and Globalization*. Cambridge: Cambridge University Press, and for 1855-1868 from Kieve, Jeffrey L. *The Electric Telegraph: A Social and Economic History*. Newton Abbot: David & Charles, 1973.

We parsed data for telegraph office location and opening date after 1870 to 1980 from a larger dataset drawn from UK Post Offices by County (<https://sites.google.com/site/ukpostofficesbycounty/>) giving the name and metadata for 37,240 UK post office locations and the starting date of telegraphic operation. After 1870 nationalisation, UK’s public telegraph network was primarily accessed through the post office network. We thank Ken Smith and Nick Bridgwater from UK Post Offices for providing these raw data.

From these raw data, we extracted, and data wrangled 14,841 UK observations of telegraph offices for the period 1870-1983. 12,527 observations for England and Wales were then pared from the full UK dataset to measure access to telegraph in those countries for which we had population data.

Accessibility data is based on the location of English and Welsh telegraph offices. Using location coordinates, we calculated telegraph coverage from 1870 based on maximum 30-minute walk access to telegraph station. This task was performed using ArcGIS. We were kindly furnished with necessary population data by the Cambridge Group for the History of Population and Social Structure.

More data for all of UK were collected from UK Parliamentary Papers and Mitchell, B. R. *British Historical Statistics,* Cambridge: Cambridge UP, 2011. We extracted number of local and national telephone calls made, and telephone stations (phone boxes) built for UK or England and Wales. Parliament’s *Annual Reports of the Postmaster General* from 1854 and related *Abstracts* 1917-1980 provided these and other data for telephone and telegraph usage. These data sources form the basis for calculating telegram volume as number of messages and data. And provided number of telephone calls and number of telephone stations.

This report accompanies the following spreadsheets containing the supplied data:

1. STATIONS\_NUMBER.xlsx
2. TELEGRAPH\_COVERAGE\_30MIN.xlsx
3. VOLUME\_ALL.xlsx
4. VOLUME\_TELEGRAPH\_BYTES.xlsx

**Value of the Data**

* These data are useful for visualising public access to the UK’s telegraph network since 1840. In addition, they give number of telephone calls made since 1898. They are important to understand the development of the UK’s historical electronic communications infrastructure before the digital age.
* These data are useful to understand when electronic communication access evolved and where in the country. It is envisaged that researchers in academia, in government, or anyone interested in the development of telegraph, telephone, or generally, electronic communication systems in the UK can benefit from this dataset.
* They might be further used for follow-up research UK inland communication. There is ample scope to combine other data, such as for rail and road connectivity, to compare and get broader understanding of trade and communication networks in the UK in the long term.
* Our raw data include Scotland and Northern Ireland telegraph offices. It would be possible to integrate the other UK nations into the dataset for telegraph office coverage and accessibility in the future at lower cost.

**Data Description**

1. STATIONS\_NUMBER.xlsx gives aggregate numbers of telephone and telegram stations by year 1839-1983 for UK. Opening and closure dates for telegram stations were collected from UK Post Offices. Cumulative sum of telegram stations operating in any year is given.
2. TELEGRAPH\_COVERAGE\_30MIN.xlsx gives annual time series 1840-1983 for coverage of telegraph measured as of telegraph offices given as % of territory in England and Wales (geographical limit of our current historical population data) within 30-minute walk of a telegraph office.
3. VOLUME\_ALL.xlsx gives annual timeseries for number of telegrams sent and telephone calls made 1870-1983 for England and Wales or UK.
4. VOLUME\_TELEGRAPH\_BYTES.xlsx gives timeseries 1840-1983 of telegraph traffic for England and Wales and UK. One character on telegraph equals 1 byte of data, following normal conventions.

**Experimental Design, Materials and Methods**

Methods are now described for each data series.

1. STATIONS\_NUMBER.xlsx

Two series give number of telegraph offices established by year. We were able to parse/extract year of telegraph office opening from 13,744 post office observations for UK post offices with telegraph services.

1. TELEGRAPH\_COVERAGE\_30MIN

This work requires more explanation. To create underlying data for our accessibility metrics it was necessary to parse a subset of data giving location and opening date of 12,559 historical telegraph offices in England and Wales, 1870-1980, and compare with population data. These data were filtered and parsed from a full dataset of 37,240 UK post office locations. For this project, a new dataset for telegraph offices was this way created for a) UK and b) England and Wales telegraph offices.

Telegraph office dataset variables parsed were:

1. Geographical address at street, town, and county levels.
2. Telegraph office first opening date.
3. Post office first opening and closing dates.

a) provided the geographical data needed to calculate accessibility from parishes in England and Wales. b) and c) allowed us to provide a time series data for the same.

Geocoding the telegraph office address variable (a) was important to create geospatial metrics for accessibility based on our 30-minute walk to telegraph office measure. UK Post Offices provided post office addresses where local public telegraph offices were located. Telegraph closure and opening dates are usually given. Post office closure dates record when telegraph offices closed down. Closure and reopen dates were collected and allowed us to make reliable calculation for number of active telegraph offices by year.

*Method of geocoding telegraph office locations:*

Available address fields for each telegraph office were concatenated, cleaned, and imported in batches of 2,500 into Google Earth Pro, the maximum number processable at once. Google Earth batch geocoding tools provided the necessary coordinates for all office addresses. Google .kml files were exported with the geocoded post/telegraph office coordinates into .csv files using an online conversion tool (https://products.aspose.app/gis/conversion/kml-to-csv) to extract decimal coordinates from .kml files to .csv files. These .csv coordinate files were then imported to Arc GIS for coverage/public access calculations.

Some post office addresses are given in an incomplete fashion in the raw data, for example with missing post codes or town names. About 10% of these addresses were ‘guessed’ by Google using its powerful database of UK place names and addresses. Google was able to locate addresses 90% with high confidence. The remaining 10% were given an approximate location. This level of precision is sufficient to calculate accessibility measures to a geographical point within 30 minutes walking distance from a given location in the vast majority of cases. There may be some error. Because there are over 12,000 post/telegraph office addresses in the full data set, we cannot practically research each individual address. We don’t expect any errors to distort results significantly because only the approximate location of the office is needed to calculate a reasonable estimate of walking times.

*Coverage metrics*

We created a pedestrian network with *ad hoc* connectors to all points representing telegraph offices in England and Wales and used this network to calculate optimal routing from each telegraph office to all parish centroids based on the assumption of a 4km/h average speed. These Origin-Destination (OD) matrices are then reduced to keep only the closer office for each centroid. We then joined these data to the point feature class produced in ArcGIS and calculated the total coverage metrics presented here.

1. VOLUME\_ALL.xlsx

Volumetric data was collected for number of telephone calls made, telegrams sent, and number of telegraph offices created UK-wide from data published in parliamentary papers and elsewhere.

1. VOLUME\_TELEGRAPH\_BYTES.xlsx

We estimated the volume of data exchanged through the network based on the average length of telegrams (20 words), and the average length of English words (4.7 words), adjusted for non-word characters, and converted this in bits based on ASCII norms. We also adjusted for the share of (longer) press telegrams sent every year.