

## **1. Fixed line telephony and the changing geographical and social structure of the UK, 1869-1985**

This exciting Ph.D. opportunity is part of Consumer Data Research Centre (CDRC) concerns with developing research resources that will make it possible for researchers to investigate how residential mobility and migration have contributed (or otherwise) to inter-generational social mobility. A project associated with CDRC research at UCL has established the extent to which most Anglo-Saxon family names were geographically localised when first coined between the 12<sup>th</sup> and 14<sup>th</sup> Centuries. Using individual records from historical Censuses, available under special licence from the (ESRC) UK Data Service, it has been possible to chart the spread of most every surname over the period 1851-1911. Research carried out by the ESRC Consumer Data Research Centre has made it possible to examine the continuation of this spread over the recent period 1997-2016, as well as the diffusion of family names imported from abroad through the UK's settlement hierarchy. What is not known are the dynamics of change over the period which postdates the release of Census data under the 100 Year Rule, and the post 1996 period in which the geographic spread of names can be inferred from digital consumer data sources. Neither is there clear understanding of whether all moves, at all time periods, left households better off, and whether the descendants of migrants or movers have benefited.

This research will use Big Data analytics and the unique data resources of the BT Archives to research these gaps in our understanding, using digitally encoded telephone directories. BT Archives holds hard copies of every telephone directory compiled between 1869 and the present day. Under an agreement with the Ancestry organisation, the directories pertaining to 1869-1985 have been scanned to high resolution and are available to amateur genealogists interested in tracing their ancestry through a subscription service. However, the data have not been used for any more aggregate analysis, and no research has been undertaken into the changing social echelons that are represented in the data. This research will rectify this anomaly by capturing names, addresses and telephone exchanges in digital form and, in the first instance, using the data to establish a continuous time series of the geographic spreads of every UK surname over the period 1851-2016.

As an end in itself, this work will equip the student with expertise in a range of quantitative skills in data capture, analysis and mapping. The student will develop an automated method of optical character recognition (OCR) data capture that accounts for the changing layout of telephone directories over time and will develop automated methods of separating entries pertaining to businesses and other organisations. It is expected that, as a minimum, the student will have sufficient time to capture data pertaining to decennial periods between 1871-1981, plus the last available year of 1985. This will facilitate comparison with available historical Census data for 1851-1911 and contemporary data for 1997-2016. Once these data have been extracted, they will be held using ESRC CDRC secure data facilities in order to minimise the risk of potentially disclosive information pertaining to individuals or households, and all reports and publications relating to the research will only contain aggregate data.

Expertise in database management systems will be developed in order to manage and georeferenced this database using address and telephone exchange information: the latter provides an interesting focus of geographic analysis in itself, given the ways in which local exchanges were established and merged over the years. This stage of the research will be

aided by data extracted from other CDRC data sources, which have georeferenced most all individual addresses recorded over the 1997-2016 period using the Ordnance Survey AddressBase Professional data product. The student will also develop and apply novel visualisation and mapping techniques, such as those used on the prototype gbnames website (<https://data.cdrc.ac.uk/gbnames/>) in order to map the contagious and hierarchical diffusion of names as implied by the raw data.

Although of interest as an end in itself, this work will also establish a baseline against which the social and geographical selectivity of telephone line rental may be gauged, given that 'opt out' from inclusion in public telephone directories did not become widespread until the advent of mass telephone marketing in the 1980s and 1990s. Decennial individual level data for the period 1871-1911, and comparison with corresponding telephone records will make it possible to understand the geography of penetration of the telephone service amongst different family groups and, along with socioeconomic data from the Censuses, the social selectivity of the service – thus, for example, charting the spread of a technology from the highest social echelons to tradesmen and other workers for whom the service was an important tool as well as a connection to social networks. ESRC CDRC research has also developed algorithms that enable the residential mobility of households to be tracked over time, and the student will adapt these in order to trace telephone subscriber moves over space. Together with numerical analysis of increases in the subscriber base over time, this will provide an important indicator of the importance of fixed line telephony as an indicator of the changing social and spatial structure of the UK.

The research will enable the student to make an important contribution to the literatures on migration, social mobility and regional development. At the same time, directory entries pertaining to business establishments and organisations will be used to identify the changing regional composition of businesses across the UK. Although providing no direct commercial benefit to BT Plc, it will demonstrate the wider economic and social value of preservation of historical records pertaining to the innovation and roll out of the major communications innovation of the late 19<sup>th</sup> and 20<sup>th</sup> Centuries.

**Applications (cover letter and CV) should be sent to Professor Paul Longley**  
[p.longley@ucl.ac.uk](mailto:p.longley@ucl.ac.uk)