

Power to the Population? The Population Census under Review

Lynn Killick, Edinburgh Napier University, Scotland
Alistair S. Duff, Edinburgh Napier University, Scotland
Mark Deakin, Edinburgh Napier University, Scotland
Hazel Hall, Edinburgh Napier University, Scotland

The Asian Conference on Technology, Information & Society 2015
Official Conference Proceedings

Abstract

Many regard the population census as the backbone of national statistics. It is also regarded as a national institution; a data source held in high regard by the academic, policymaker, historian and genealogist alike. However, technological advances, pressure on resources and the availability of alternative information about the demography of the population have led to a recent review of the census. The results of this review have led the UK to develop a population census that in future will be conducted online and augmented by data derived from other government sources, for example, health and social care records, without the need for explicit consent of the population. It is, therefore, an opportune time to build upon previous studies relating to privacy and the census and examine the impact that these confirmed changes to the population census in the UK will have on public perceptions of the confidential nature of the census, with a particular focus on information security, privacy and ethics. This paper will discuss the results of a study examining public attitudes to an online census, information sharing between government agencies without explicit consent, and attitudes to private companies processing census data. The census office have made changes to the census have with limited input from the public and lead us to question where ultimate power lies? Is it with those making the changes, or those providing information, to the census?

Keywords: Privacy, confidentiality, population census, information sharing, information society

Acknowledgements: this work is supported by the Arts and Humanities Research Council (grant number AH/K002899/1).

iafor

The International Academic Forum

www.iafor.org

Introduction

History tells us that those in power from Babylonian times to the present day have understood the importance of knowing their population (Grajalez, Magnello, Woods, & Champkin, 2013) (Missakoulis 2010). It is, however, the case that the rationale for the count, its method and the subjects of the count, have changed dramatically from the early counts conducted in Babylonian times. The question is, as the nature and purpose of the count have changed over time, has the power shifted from the state to the population?

In an attempt to answer this question I would like to share with you the preliminary results of some work that I have undertaken to assess the perception of the public and policy-makers to upcoming changes to the population census in Scotland. However, I think it is important first to look at the origins of the census and, in particular, the position of power in the population census.

The census and the state

In early history, the censuses were predominantly taken to assess the wealth of a nation, or perhaps more specifically a nation's ruler, by establishing the assets at their disposal. For example, the count of the population enabled the Egyptians to establish the workforce available to them to construct the pyramids. The Romans are also well documented for their use of the census to ascertain taxes due and the availability of conscripts for the Army. Literature examining historical censuses suggest that in the main, censuses were predominantly undertaken to count men available to till the land, fight in wars and importantly pay taxes based on land and livestock owned (Coleman, 2013, Grajalez et al., 2013). The 17th & 18th centuries, however, brought a change in focus for the census. Governments came to understand that information about the population of their nation could be used as evidence to inform policy decisions, with specific regard to policy decisions that could improve health and social outcomes for the population (White, 2011).

To obtain this information the state enters into a contract with the population. The population are compelled to complete the census within a set timeframe, and the state provides a commitment to hold census information securely and importantly to ensure published census results do not identify individuals or households. (In the UK census records must be kept confidential for 100 years)

However, despite a commitment to security and confidentiality past studies have identified an abuse of power on the part of the state. For example, US Census information was accessed to locate people living in some US states and identifying as Japanese for internment (Anderson & Seltzer, 2009). Perhaps more notably, the Nazi party abused the confidentiality and purpose of the census to use census information to locate individuals who could help or hinder its vision for an Arian society (Aly & Roth, 2004). So history tells us that the census is indeed a powerful tool for the government of the day, a tool that has the potential to be used or abused.

There are of course no reported instances of data abuses in the UK, nor is there any evidence that confidential information originating from the census is shared between government agencies. However, the public may be justified in having misgivings regarding the security of personal information collected by the census. The most recent privacy impact assessment for the census highlights the confidential nature of the census and the robust data control measures in place. It also suggests, by

reference to the Statistics and Registration Service Act 2007, that the census data could be used for alternative purposes as Section 39 (4)(g) provides an exception, allowing information to be released, if it is in the interests of national security, to an intelligence agency.

The census in Scotland has at its core a commitment, and legal obligation, to ensure the right to confidentiality of those who participate in the census. There is, of course, a risk that any endeavour that seeks to capture and publish detailed information obtained from individuals could lead to data breaches.

The census and the population

Research by Heeney (2012) explains that the potential for data breaches is one factor that may affect public participation in the census; another is the intended use of the data provided in the census. Furthermore, research by Cullen (2008) highlights concerns regarding data handling, particularly by private organisations, are a factor that has the potential to affect the trust of the public in the census adversely.

In the UK, there are powerful examples of the public withdrawing from the census to great effect. The withdrawal from the census by suffragettes in the 1911 census is well documented (White, 2011); in this instance a very public protest to raise awareness of the campaign to grant all women a vote. More recently swathes of people avoided the count in 1991 in rebellion against the poll tax (Simson and Dorling 2004), in this instance, a quieter, although sizeable, avoidance of the count driven by concerns that census data would be shared without restriction to other government departments. The impact of the ‘missing millions’ (Simpson et al. 2004) was not fully appreciated until after the count at the point of distribution of funds from central to local government, and it is only by comparing the 2011 census with the 1991 census that statisticians were able to identify that the majority of the missing millions in the 1991 were young adult males. More recently across the UK we see significant numbers allegedly misrepresenting themselves through their declaration of religious affiliation with the Church of the Jedi Knight.

Brunton & Nissenbaum (2011) provide a political theory of obfuscation, they define obfuscation in the context of data gather as ‘producing misleading, false, or ambiguous data to make data gathering less reliable and, therefore, less valuable.’ In the context of the census, perhaps with the exception of the suffragette protest, it is not clear that the provision of false or ambiguous data is to make the census less valuable. Rather recent studies (See Heeney 2012 & Cullen 2008) suggests a decision not to participate in the census is more likely in an attempt protect privacy due to a lack of trust regarding data confidentiality. My work seeks to explore the potential for obfuscation in more detail. To understand the motivations of individuals and the risks to the census it is necessary to understand more about the planned changes to the Scottish census.

The Future of the Census

Firstly, the process for data collection in the census is set to change. The next census in Scotland will be conducted predominantly online. Online completion was an available option during the 2011 census. The majority of census returns (80%) were returned using the traditional form. The relatively low take up reflects the soft marketing approach employed by the census office. The online response rate of 20% was considered a success by the census office. It produced returns in a manageable

number that the systems were able to cope with. Also, the online returns provided better data quality in comparison with handwritten returns, with fewer errors and increased efficiencies in data processing (Dr J Goodlet-Rowley, Seminar, December 2013).

In addition to a predominantly online the census, the census office has advised that census results will be augmented by information obtained from administrative records. Administrative records contain identifiable information gathered for administrative activities as opposed to statistical endeavours. Such movements of data may be legal. However, ethical issues are present. The use of administrative records raises issues of informed consent and to an extent the 'right to be forgotten' as highlighted by Ausloos, (2012). For example, information provided to register with a doctor or tax a vehicle is not provided with explicit consent to be retained ad infinitum or shared with external agencies.

The act of the census itself, however, is found to be compliant with Article 8 of the Human Rights Act as the process of obtaining information by compulsion is deemed necessary to support the aims of a democratic society.

The census has been described as a public institution (Anderson 2008, p2). The expectation may then follow that the sensitive process of collecting and publishing census information would be a public endeavour, operated in the public sector. However, recent census exercises, for example across the UK and also New Zealand, have outsourced elements of the census. This outsourcing, or perhaps more specifically, the choice of contractors, has not been without controversy.

To illustrate the importance of securing public trust Cullen (2008) described the careful planning that led to the success of the 2006 New Zealand's first census allowing online responses. Factors such as perceptions of information security were paramount, particularly as a private sector company was being used to collect the data.

The Joseph Rowntree Reform Trust Ltd in a review of government databases, including the English and Welsh census, noted the use of a private sector company, Lockheed Martin, as an information security risk. Public scrutiny relating to the award of Lockheed Martin for the English and Welsh 2011 census and a comparable firm, CACI, for the Scottish 2011 census, resulted in a review of the contract awarded to these firms. The data processing rights were reviewed and these firms, both with links to foreign intelligence agencies, were prevented from storing the census data. It is likely that the future Scottish 2021 census will involve a considerable procurement exercise to ensure there is appropriate infrastructure in place to support a predominantly online census capable of capturing the information relating to over five million people over a short time frame. It is not clear if procurement legislation will prevent companies with involvement in intelligence operations from bidding for the delivery contract.

Once collected, census data is released in the years following the census in an anonymised format and released in its entirety 100 years after the census has taken place. Information generated from the census is open data. It is possible for any individual or organisation to download data tables or access a variety of tools available on the census website. Also, individuals and organisations can make direct requests for particular results from the census.

(<http://www.scrol.gov.uk/scrol/help/scrolHelp.jsp>)

The open data generated from the Census can be accessed by anyone and linked to any other data source. Palmer (2013) highlights how Australian census information, combined with anonymised customer records of an Australian bank, to create a comparison website, provided the conditions to identify individuals. This suggests that it may be possible to identify individuals if statistical data from the Scottish census is combined with data gleaned from other sources..

This notion that information given in good faith for one purpose could be shared and use for another, divides opinion, illustrated for example by comments regarding the perceived security of census information at

http://www.theregister.co.uk/2011/03/25/confidentiality_of_census_data_not_guaranteed/.

Some comments suggest anger and deep distrust in the census due to the possibility of sharing information with intelligence agencies; others remark that the Facebook generation should be accepting that there is no such thing as private information. An online approach to the census may be attractive due to improvements made to the quality of the data obtained, (Baffour et al. 2013, Coleman 2013, Cullen 2008, Deonandan 2013) and the cost-savings generated as a result of reduced data cleansing and more efficient data processing. An online future does, however, present challenges both for data users and the public as well as those responsible for administering the census. Issues of access to technology will hinder many individuals participating easily in the census. Language barriers, availability of hardware, confidence and disabilities can all impact on an individual's ability to participate online. In addition, a connection with the census may be lost if it is reduced from its current historical record featuring ancestors handwritten entries to an online box check survey. This sense of value in the census may be further diminished if individuals cannot see an explicit link between the census and policy decisions to their benefit, or the benefit of wider society. Concerns regarding data privacy, trust in government and the perceived low risk of sanction for non-compliance are other factors that may also result in individuals deciding not to participate in a future census if it transfers to a predominantly online activity.

The studies to date do raise issues of trust and concerns regarding data security. However, the recent studies looking specifically at attitudes to the census, and specifically an online census, have been small in scale, have not specifically sought the views of Scotland's population and, therefore, call for more attention. Also, it is not yet clear from the literature, the extent to which the population in the UK are aware of issues such as data breaches in other censuses and to what extent this may affect their participation in either an online census or if they would be opposed to data sharing across government agencies.

This study

As part of a wider doctoral study, I am investigating the perceptions of the public and policy-makers to the census. In this paper, I discuss findings generated from semi-structured interviews with policy-makers who may work with census information and emerging findings from an online survey created to gather responses from the public. The interviews were conducted in 2015 with 20 policy-makers working in Government (both local and national), Non-departmental Government bodies (such as Health Boards) and with individuals holding policy roles in the third sector and those working in governance roles as either non-executive directors of public bodies or regulators in organisations set up by parliament but independent of government.

The online survey was designed using Novisurvey and was publicised using social networking platforms such as LinkedIn and Twitter. The survey was aimed at people who either work or are resident in Scotland. The results shared in this paper represent complete returns from 124 individuals, all of whom reported a Scottish connection through work or residency.

The information presented here relates specifically to questions regarding perceptions regarding data security and data sharing set against an exploration the respondents' view of the purpose of the census.

The online survey results suggest that the place of the census as a public institution (Anderson 2008) is still relevant as 89% of survey respondents believed that taking part in the population census was an important part of citizenship. Furthermore, 53% reported that the purpose of the census was to develop public policy.

This understanding of the place and purpose of the census is echoed in the semi-structured interviews with policy makers who reported the view that the census was an important societal asset albeit one that could be applied more effectively to the policy-making process (Killick, Hall, Duff & Deakin, in press)

Policy-makers expressed concern about the online move, although issues of data security or concerns regarding procurement are not expressed as priority issues. Rather, issues relating to broadband availability in remote and rural areas coupled with low information literacy skills and limited access to Internet-enabled devices are mentioned. The concerns here relate to under-enumeration as a result of Scotland's patchy digital infrastructure.

Perhaps reflecting the online nature of the survey 76% of respondents indicated that they would prefer to complete their next census online with 55% of all respondents indicated that the thought that completion of the census online offered a secure method. This suggests that the convenience of Internet-completion outweighs concerns regarding data security.

Considering the procurement of the census, echoing the findings of Heeney (2012), the survey respondents were not comfortable with private sector companies handling census data. When asked specifically about UK-based private sector firms 52% stated that they were not comfortable with such firms handling their data, with only 26% confirming they would be comfortable with this situation. When asked regarding their attitudes regarding non-UK based firms handling the data there was a marked shift with only 0.05% (6 people) confirming they would be happy for an overseas firm to handle their data and 81% stating that they would not be comfortable with overseas firms handling their census data. The strength of objection to private sector firms being involved in the handling of data suggests that the census office will need to undertake activity to alleviate concerns, this finding is consistent with the work of Cullen (2008).

When asked about plans to augment census data with information gleaned from administrative records policy-makers working in institutions external to government (third sector and regulatory bodies) voiced concern both about the quality of the information to be input to the census and also expressed caution that such 'augmentation' may be the first step to replacing the census. The specific concern being that while a variety of sources can provide a population count, only the census can provide a reliable dataset that reports on the diversity of the population.

Survey respondents were in the main happy to provide information to the census (62%), albeit noting that legal obligation to complete a return was a driving factor in

their participation (47%). The attitudes, however, were not as positive regarding information sharing, with 45% comfortable with the information they had provided to other government departments being shared with the census and 34% noting an objection.

The survey respondents were also asked for their opinion regarding information being shared from the census to other government departments. First they were asked if they were comfortable with anonymous information from the census being shared, first with government departments such as HMRC (Tax) or Health & Social Care.

The vast majority of respondents noted that they were comfortable with such information sharing with only 16% expressing that they were not content. Similarly, the same question regarding anonymous information sharing but in this instance with government security agencies generated a similar response again with 16% of respondents expressing dissatisfaction. However, when asked regarding their attitudes regarding identifiable information being shared with other government agencies, there is a marked difference with 60% reporting opposition to identifiable information being shared with any government agencies. These results suggest that the confidential nature of the census is important to members of the population.

Another factor may be levels of trust, 66% of respondents noted that they trusted the census office to treat their personal information with respect for their privacy, compared to a lower number (58%) reporting a similar level of trust in 'government'.

Conclusion

These initial findings suggest that if the next census is to be successful those responsible for its administration must balance the necessary infrastructure challenges to ensure effective enumeration of those who may be digitally excluded, with activity to ensure the population has confidence in an online census and importantly the confidential nature of the census and its outputs. A failure to do so has the potential to alienate sections of the population and may lead to a repeat of the 1991 census.

The initial results of this study suggest that any motivation to obfuscate that may manifest itself is likely to originate from a desire to protect one's privacy rather a deliberate desire to devalue the census. More work would be welcomed in this area particularly as if there is a significant undercount in the 2021 census it could undermine the future of censuses beyond 2021.

References

Aly, G, Roth, K H, (2004) *The Nazi Census: Identification and Control in the Third Reich*. Philadelphia: Temple University Press

Anderson, M (1988) *The American census: A social history*. New Haven: Yale University Press

Anderson, M (2008) The Census, Audiences, and Publics. *Social Science History*, Volume 32 (1), p1-18

Ausloos, J (2012) The 'Right to be Forgotten' - Worth remembering? (2012) *Computer Law & Security Review*, 28 (2), p143-152

Balarajan, R., & Raleigh, V. S. (1994, 1994/07/30/). Public health and the 1991 census: non-random underenumeration complicates interpretation. *British Medical Journal*, 309, 287+.

Coleman, D. (2013). The Twilight of the Census. *Population and Development Review*, 38, 334-351.

Cullen, R (2008) New Zealand's 2006 Census Online: A Case Study. *Integrated Series in Information Systems*, 17, p647-670

Cullen, R (2009) Culture, identity and information privacy in the age of digital government. *Online Information Review*, 33 (3), p405-421

Grajalez, C. G., Magnello, E., Woods, R., & Champkin, J. (2013). Great moments in statistics. *Significance*, 10(6), 21-28.

Heeney, C. (2012). Breaching the Contract? Privacy and the UK Census. *Information Society*, 28(5), 316-328.

Missiakoulis, S. (2010). Cecrops, King of Athens: the First (?) Recorded Population Census in History. *International Statistical Review*, 78(3), 413-418.

White, I. (2009) The 2011 Census taking shape: methodological and technological developments. *Population Trends*, 136, p64-72, Palgrave Macmillan

White, I (2010) No vote – no census: an account of some of the events of 1910-1911. *Population Trends* 142, Office for National Statistics

Simpson, S., & Dorling, D. (1994). Those Missing Millions: Implications for Social Statistics of Non-response to the 1991 Census. *Journal of Social Policy*, 23(04), 543-567