



Department of
**Finance and
Personnel**
www.dfpni.gov.uk

Online Data Collection in the Census

October 2014



The Northern Ireland Statistics and Research Agency

The Northern Ireland Statistics and Research Agency (NISRA) is an Executive Agency within the Department of Finance and Personnel (DFP) and has been in existence since April 1996. The Agency also incorporates the General Register Office (GRO) for Northern Ireland. NISRA's core purpose is to provide a high quality, cost effective, statistics, research and registration service that informs policy making and the democratic process and the wider public.

The overall corporate aims of NISRA are to:

- provide a statistical and research service to support decision making by Northern Ireland Ministers and Departments and to inform elected representatives and the wider community through the dissemination of reliable official statistics; and
- administer the marriage laws and to provide a system for the civil registration of births, marriages, civil partnerships, adoptions and deaths in Northern Ireland.

NISRA can be found on the internet at www.nisra.gov.uk

Northern Ireland Statistics and Research Agency

McAuley House
2-14 Castle Street
Belfast
BT1 1SA

Contents

1.0 Introduction 3

2.0 Drivers for Change 4

3.0 The Beyond 2011 Programme 5

4.0 Electronic Interaction with Government 6

5.0 Internet Access in the UK and Beyond 9

6.0 The Internet as a Data Collection Mode in the Census 12

7.0 Benefits of an Internet Data Collection Mode 15

8.0 Suggested Design Artefacts for an Online Data Collection Mode 20

9.0 Conclusions..... 24

1.0 Introduction

The Census is the largest statistical exercise undertaken in Northern Ireland and provides comprehensive statistical data covering many aspects of life, such as population distribution and demographic, social and economic characteristics.

The Census like any other statistical investigation amongst the population is subject to change with new questions and modes of collection. This paper outlines the move towards online data collection and the benefits this may bring.

The objective of this paper is twofold. Firstly to stimulate discussion and secondly to start the move towards a way forward with respect to an online data collection mode for the Census.

Section 2 of this report outlines the drivers for change in the Census and Government's wider digital agenda. Section 3 of the report introduces the Beyond 2011 programme. Section 4 reports on research by OFCOM into electronic interaction with Government. Section 5 examines internet access in Northern Ireland, the UK and beyond. Section 6 examines the use of the internet as a Census data collection mode across the UK and beyond. Section 7 summarises the benefits of an internet response mode by examining the experiences of a selection of countries. Section 8 suggests some design artefacts to assist in developing an online mode and increasing internet responses. Section 9 comments on the use of the internet as a data collection mode for the wider Northern Ireland Statistics and Research Agency (NISRA) survey programme and section 10 draws some initial conclusions.

2.0 Drivers for Change

Rising costs and an increasing population, combined with a period of austerity in the UK has increased the impetus to reduce costs across Government. As with most statistical investigations amongst the population, fieldwork comprises a major element of the cost of the decennial Census. Data collection by the internet is hoped to produce cost savings without any unacceptable reduction in the quality of the Census results. The population expects Government to use technology for the efficient and effective delivery of public services. The political impetus is in this direction and the expectation is that public interaction with Government will be “digital by default”¹. In relation to NISRA, the Northern Ireland Government’s statistical agency, the Census is a key area of interaction with the public.

Governments, both locally and nationally, recognise that not all citizens can utilise digital technologies. In order to help these digitally excluded people Government has a range of strategies and initiatives some of which are described below. The UK Government has an overarching Digital Strategyⁱ that includes an action to reduce digital exclusion. To achieve this aim a Digital Inclusion Strategyⁱⁱ was launched in April 2014. In Northern Ireland the Digital Inclusion Unit is based within the Delivery and Innovation Division in the Department of Finance and Personnel. The unit is responsible for digital inclusion, and focuses on ensuring that the people of Northern Ireland are capable of accessing new digital services and establishing, funding and operating a number of digital support projects.

Launched on the 15th April 2014, Go ON NIⁱⁱⁱ aims to increase the basic online skills of people, small business and charities in Northern Ireland. The programme, in partnership with Go ON UK, aims to deliver a 25 per cent reduction in people below the basic online skills threshold in 12 months. The programme’s website reports:

“Over 345,000 adults in Northern Ireland – just under a quarter – lack the Basic Online Skills needed to send and receive email, use a search engine, browse the internet and complete online forms. Almost a third of businesses and charities are also missing out”.

Reducing the levels of digital exclusion will assist in maximising online responses in the Census.

The Northern Ireland administration is also proactive in the area of online service delivery. The Digital Transformation Service – Citizen Contact Strategy^{iv} looks at how citizen-facing services are delivered now, and sets out a vision for the future. One of the six main principles is to:

- *“Promote a ‘digital first’ ethos, meaning that digital online services are the primary means of interacting with citizens and businesses”.*

The Citizen Contact Strategy which promotes a “digital first ethos” implies Government support for NISRA to introduce an online option for the Census.

3.0 The Beyond 2011 Programme

Beyond 2011 is the programme being undertaken by the Office for National Statistics (ONS) to assess the future for the Census and population statistics in England and Wales. Parallel, but proportionate, projects are being undertaken by the National Records for Scotland (NRS) and NISRA.

The Beyond 2011 programme was tasked with reviewing approaches to the Census including possibilities to modernise the Census process while considering the issues of cost, benefit, statistical quality and public acceptability. The programme’s first major milestone in England and Wales was the National Statistician’s^v recommendation to Government for the 2021 Census in England and Wales.

The National Statistician recommended a predominantly online Census in 2021, which should produce a cost saving, and the use of administrative data and surveys to “*enhance*” Census statistics.

The Census options for England and Wales were reviewed independently, prior to the National Statistician’s recommendation, by Professor Chris Skinner^{vi} (London School of Economics). In relation to the internet option Professor Skinner and his team concluded:

- *“ONS has considerable experience with traditional census methodology and has growing experience of the use of the internet for data collection. We*

believe it has a sound appreciation of the methodological challenges posed by this option”.

Skinner also noted that an online Census:

- *“Would represent a natural evolution of the traditional census, drawing on technological innovations and developments in best practice for census taking around the world”.*

Skinner in reaching his conclusion set two provisos. Firstly all cases where an online response is not received will be followed up fully. Secondly that prior to implementation there would be thorough research into possible mode effects.

Independently the NRS^{vii} announced that the 2021 Census in Scotland will, as in England and Wales, be predominantly online.

The most recent Census to be carried out in the Republic of Ireland (RoI) was Census 2011. Unlike in the UK there was no option to respond online. The next RoI Census will take place on Sunday 24th April 2016. The methodology to be employed in 2016 was described by the National Statistics Board as *“no change”^{viii}*, in other words the 2016 Irish Census will follow the same methodology as in 2011 and no online response mode will be available.

The approach to Census 2021 in the UK has a strong focus on the internet. This approach sits within the Government’s wider digital agenda set out in section two of this report.

4.0 Electronic Interaction with Government

People in Northern Ireland regularly purchase goods and services online.

Government processes such as annual tax returns and applying for a driving license can now be done online. Research by OFCOM^{ix} into adults’ media use and attitudes examined how people interacted with Government. OFCOM reported that in 2013, 61 per cent of UK internet users complete Government processes online (via email/website). This made it the preferred option for this activity. Respondents were asked about their reasons for completing Government processes online. For those adults that have completed a Government process online convenience was the most

popular reason (84 per cent). Four in ten (40 per cent) said it was quicker than in person, either face-to-face or on the phone. One in eight (13 per cent) of internet users who had not completed a Government process online said that it is because they were unaware that it was possible.

The Continuous Household Survey (CHS) 2012/13 in Northern Ireland reported that 25 per cent of internet users accessed Government services online. Even allowing for differences in the definition of completing Government processes (OFCOM) and accessing Government services (CHS) internet users in Northern Ireland appear less likely to interact electronically with Government than UK internet users. This finding implies that the objective to have a predominantly online Census in 2021 may be more difficult in Northern Ireland than in the other countries of the UK.

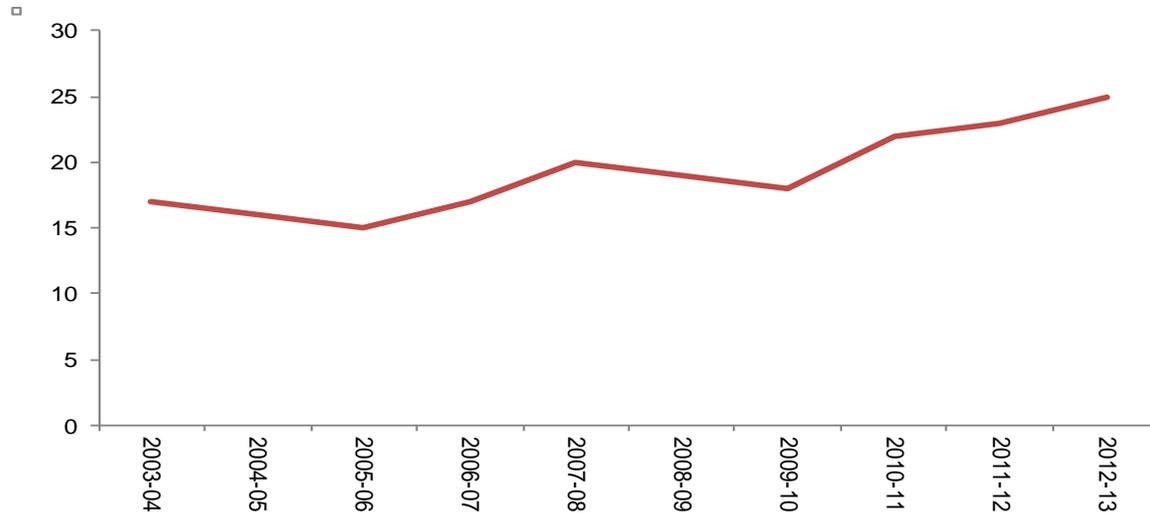
The most common reason^x for not accessing the internet reported in CHS 2012/13 was a lack of interest, cited by 46 per cent of adults. Forty-three per cent of adults not accessing the internet did not have a computer, 23 per cent lacked confidence to do so and one fifth felt too old to access the internet. Although the number of adults not accessing the internet has decreased over time the proportion citing age as the reason increased from 14 per cent in 2003/04 to 21 per cent in 2012/13 indicating a residual group of people who perceive their age to be a barrier to accessing the internet.

Although internet users in Northern Ireland are less likely to interact electronically with Government than those across the UK as a whole the proportion doing so has increased. Figure 1 shows that in 2003/04 17 per cent of Northern Ireland internet users accessed Government services online. By 2012/13 this had increased by eight percentage points (17 per cent to 25 per cent) with the increase concentrated in later years.

OFCOM also questioned respondents about how they preferred to interact electronically with Government in 2013. The majority of internet users who had ever completed a Government process online prefer to use a PC/laptop/netbook (83 per cent). This preference was more prevalent among those aged over 45 years and least prevalent among those aged 25-34 years. More modern technologies were less

evident. One in ten (nine per cent) use a tablet and four per cent use a smartphone to complete one or more Government processes online.

Figure 1 Proportion of Internet Users Accessing Government Services Online in Northern Ireland.



Source Continuous Household Survey NI.

The analysis presented so far in this section of the report suggests that promotion, which focuses on the convenience and speed of use of the internet to complete Government processes, would be beneficial in promoting electronic interaction with Government. While PC / laptop / notebook are the main devices currently used for internet access, this may not be the case by 2021.

OFCOM reported on the take-up of electronic devices generally in 2013. More UK adults, especially older adults, are now going online, using a range of devices. There has been a nine percentage point increase in those aged 65+ ever going online (42 per cent: 2013, 33 per cent: 2012). However older people are still more likely to be digitally excluded than younger people. Nearly all those aged 16-34 are online (98 per cent) compared to 42 per cent of those aged 65+.

The number of adults using tablets to go online has almost doubled, from 16 per cent in 2012 to 30 per cent in 2013. While almost all age-groups are more likely than previously to use tablets, use by those aged 35-64 has doubled, while use by 65-74 has trebled albeit from a low base of 5 per cent to 17 per cent.

Six in ten UK adults (62 per cent) now use a smartphone, an increase from 54 per cent in 2012. Older adults aged 65-74 are almost twice as likely to use a smartphone now compared to 2012 (20 per cent: 2013, 12 per cent: 2012).

5.0 Internet Access in the UK and Beyond

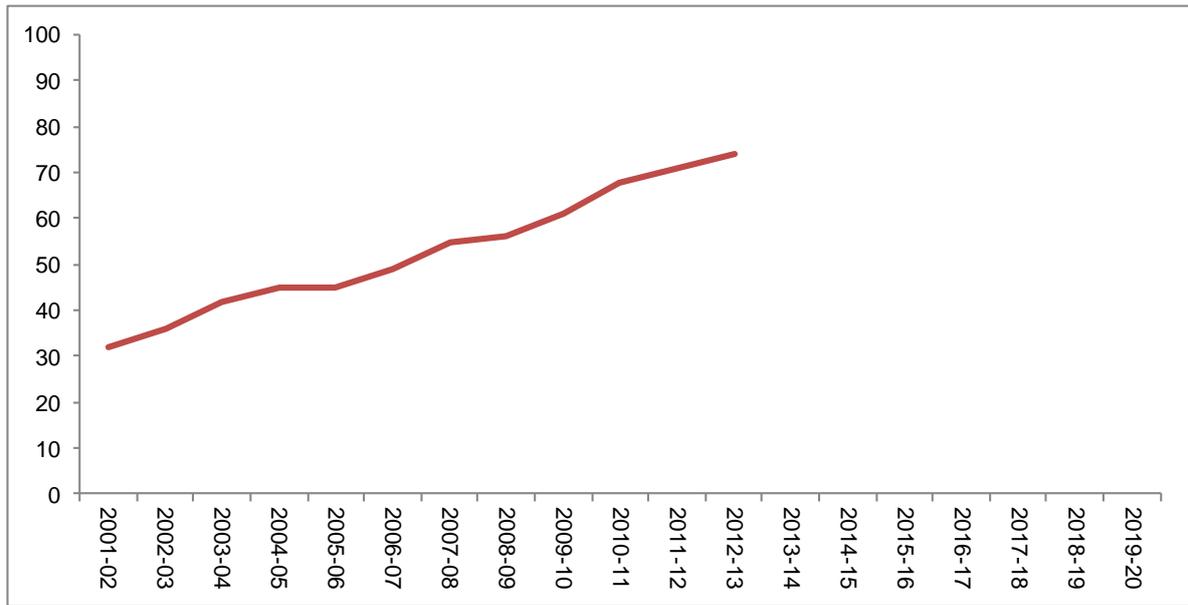
Without suitable coverage and take up of digital services, including broadband, the aim to make Census 2021 predominantly online may not be realised. This section of the report examines digital access across the UK and beyond.

Figure 2 shows the proportion of households in Northern Ireland that have internet access. Approximately three quarters (74 per cent) of Northern Ireland households have internet access in 2012/13 as reported in the CHS.

The level of internet access has grown steadily over time as shown in Figure 2. A crude extrapolation suggests that by 2021 internet access will have reached near saturation coverage. However it is unlikely that every household in Northern Ireland will have internet access at any point in time. Some will simply not want and some may not be able to afford internet access. A further group of the digitally excluded may not have the ability to interact successfully online. There will always be a residual number of households for which other means of responding to the Census or surveys are required. The design of the Census would therefore be multi-mode with respect to response options with an emphasis on an online mode.

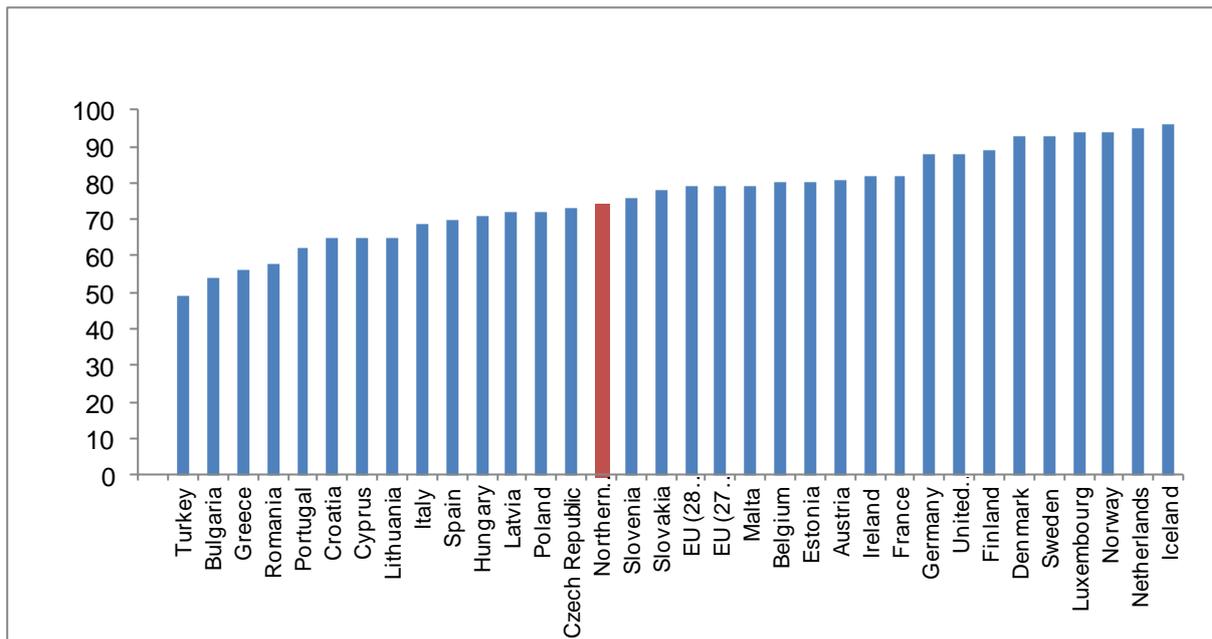
The level of internet access in Northern Ireland at 74 per cent of households is below the EU area average (79 per cent) and the level for the UK as a whole (88 per cent). Countries across Europe show a large range in levels of internet access, from a high of 96 per cent in Iceland to a low of 49 per cent in Turkey. Although below average in Northern Ireland Figure 2 suggests a sufficiently high level of coverage by 2021 to provide an opportunity to develop the internet as a data collection mode.

Figure 2 Household Internet Access in Northern Ireland.



Source; NISRA Continuous Household Survey.

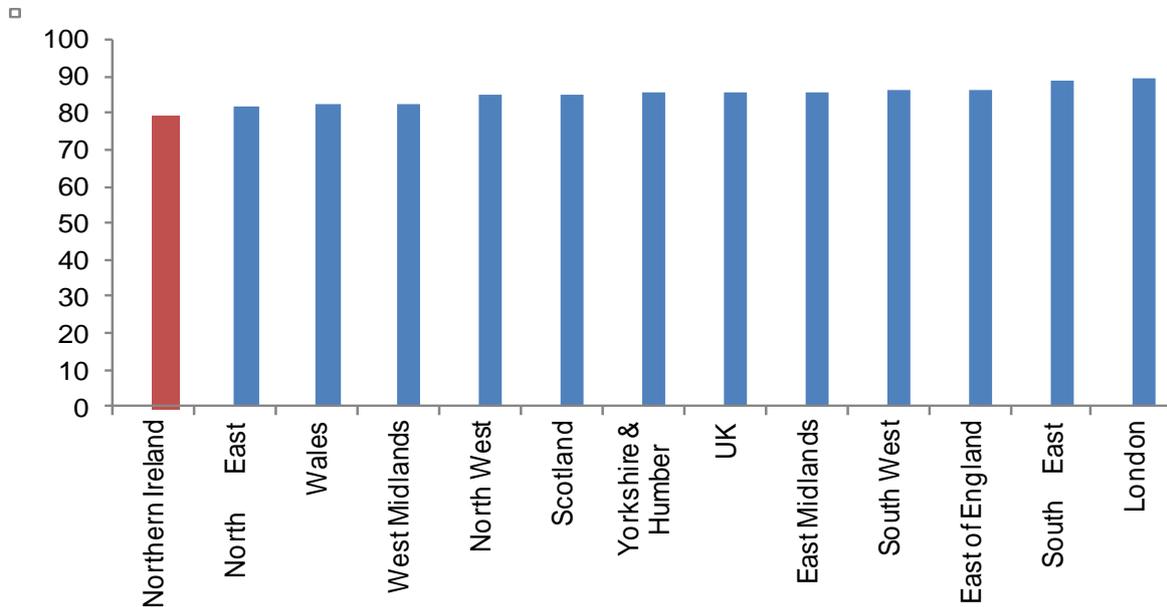
Figure 3 Percentage of Households with Internet Access in 2013; EU Area.



Source Northern Ireland Continuous Household Survey, all other data Eurostat

Figure 4 compares internet access in the first quarter of 2013, across UK regions. Northern Ireland has the lowest level of internet access across UK regions.

Figure 4 Internet Access Across UK Regions Quarter 1 2013.



Source Office for National Statistics

Broadband access is lower than overall internet access. In 2012/13, 70 per cent of households in Northern Ireland have broadband. This level of broadband access is below that in GB where 83 per cent of households have broadband. Across the EU area over three quarters of households (76 per cent) have broadband access, ranging from a high of 95 per cent in Iceland to a low of 46 per cent in Turkey.

Northern Ireland performs poorly in comparison to EU averages and other countries in the UK in respect of the take-up of digital services. This would indicate that to make Census 2021 predominantly online may be more difficult in Northern Ireland than in the other countries of the UK.

6.0 The Internet as a Data Collection Mode in the Census

Traditionally people could reply to the Census through either self-enumeration by a paper questionnaire or face-to-face with a Census enumerator. The introduction of an online mode is a further method of self-enumeration.

The 2010 round of population and housing censuses witnessed a rise in the use of the internet for data collection. Of the 138 countries or areas that responded to the United Nations^{xi}, 32 (23 per cent) utilised the internet for data collection

Table 1 shows the proportion of Census returns made online for a selection of countries. The take-up of the internet response mode is shown for the first Census in which it was offered regardless of the actual year in which the Census took place.

In 2011, 15 per cent of Northern Ireland Census returns were via the internet. At the time 71 per cent of households in Northern Ireland had internet access. The large differential between online responses and the level of internet access indicates that there is significant potential to increase online Census returns in 2021.

Table 1 Proportion of Census Returns via the Internet First Census in Which an Online Response Mode was Offered.

Country	Percentage of Census Returns made via the Internet – First Census to allow online return	Date of Census
Northern Ireland	15	2011
England & Wales	16	2011
Canada	18	2006
Australia	9	2006
New Zealand	7	2006
Kuwait	15	2011
Portugal	50	2011
Singapore	15	2000

Sources: Statistics Canada; Office for National Statistics; Northern Ireland Statistics and Research Agency; Australian Census Bureau; Statistics New Zealand; Central Statistics Bureau State of Kuwait; Statistics Portugal and US Census Bureau; Department of Statistics Singapore.

Northern Ireland compares well with other countries offering the internet mode in the Census for the first time, with the exception of Portugal. The success of the online mode in Portugal is an anomaly, indeed Statistics Portugal^{xii} was “surprised” by the level of internet responses. Statistics Portugal attributed this result to a remuneration

Online Data Collection in the Census

model that did not distinguish between response modes. As online responses required less resource from the enumerator, they encouraged people to answer via this mode.

The level of internet response in the Singapore Census 2000, at 15 per cent, is noteworthy when considered against a backdrop of only 32 per cent of the population being internet users at the time^{xiii}.

In 2000, the U.S. Census Bureau established itself as a digital pioneer by making the United States one of the first countries to use an internet-based data collection method for its population census. The online mode was trialed in 2000 but was purposely not widely publicised and consequently the level of response was low 0.07 per cent (approximately 71,000 responses). Unfortunately, citing the potential risk that personal data could be hacked, as well as logistical concerns and lack of proof that an online option would save money or improve the response rate the Census Bureau cancelled plans to use the internet for data collection in the 2010 population census. The Census Bureau decision not to have an online mode in 2010 was made even though results of an internet based response mode were positive during the Census dress rehearsal in 2005. The American Community Survey (ACS) which supplements the Census short form currently has an online response mode. Figure 3a shows the portal for completing the ACS and Figure 3b the person based household screen from the 2005 Census test. The short form Census will have an online data collection mode in the 2020 Census round to complement the online mode for the ACS. The Census Bureau estimate that half of respondents that would have replied by mail will respond online to the 2020 Census.

As with many national statistics institutes, including NISRA, the U.S. Census Bureau conducts a wide array of other surveys in addition to the decennial population census. Among these are the Current Population Survey, the American Housing Survey, the Census of Governments, the Annual Survey of Manufactures, and the Economic Census. Since 2002, businesses participating in the Economic Census have been permitted to submit their responses over the internet.

Online Data Collection in the Census

Figure 3a The Portal for Online Response to the American Community Survey.

United States
Census
Bureau

U.S. Department of Commerce | Economic and Statistics Administration

American Community Survey

Welcome to the American Community Survey. You will need the materials we mailed to you to start the survey.
All the information that you provide will remain completely confidential.

[Para completar en español, oprima aquí.](#)

Please Log In

ACS 999 999 999 99 999 9999 99
SEQ999-99999

#####

Example User ID

Enter the 10-digit User ID found below the barcode on the materials we mailed to you.

User ID: -

Login

The Census Bureau estimates that, for the average household, this survey will take 40 minutes to complete, including the time for reviewing the instructions and answers. You may email comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing burden, to: Paperwork@census.gov. Use "Paperwork Project 0607-0810" as the subject. Or you may send comments to: Paperwork Project 0607-0810, U.S. Census Bureau, 4600 Silver Hill Road, AMSD - 3K138, Washington, D.C. 20233.

Respondents are not required to respond to any information collection unless it displays a valid approval number from the Office of Management and Budget. The 8-digit number appears in the left side of the green bar at the bottom of the survey screen.

Source US Census Bureau

Figure 3b Person-based Household Screen: US Census Rehearsal 2005.

U.S. Census Bureau

2005 National Census Test

Household Persons Review/Submit

Help | Privacy Policy

Before you answer the first question, count the people living at your house, apartment, or mobile home using our guidelines.

- We want to count people where they usually live and sleep.
- For people with more than one place to live, this is the place where they sleep most of the time.

[EXCLUDE these people](#) [INCLUDE these people](#)
We will count them at the other place.

How many people were living or staying in this house, apartment or mobile home on September 15, 2005?

people

Were there any additional people staying here September 15, 2005 that you did NOT include in the question above?

No Yes

- Children, such as newborn babies or foster children
- Relatives, such as adult children, cousins, or in-laws
- Nonrelatives, such as roommates or live-in baby sitters
- People staying here temporarily

Is this house, apartment, or mobile home -

- Owned by you or someone in this household with a mortgage or loan?
- Owned by you or someone in this household free and clear (without a mortgage or loan)?
- Rented for cash rent?
- Occupied without payment of cash rent?

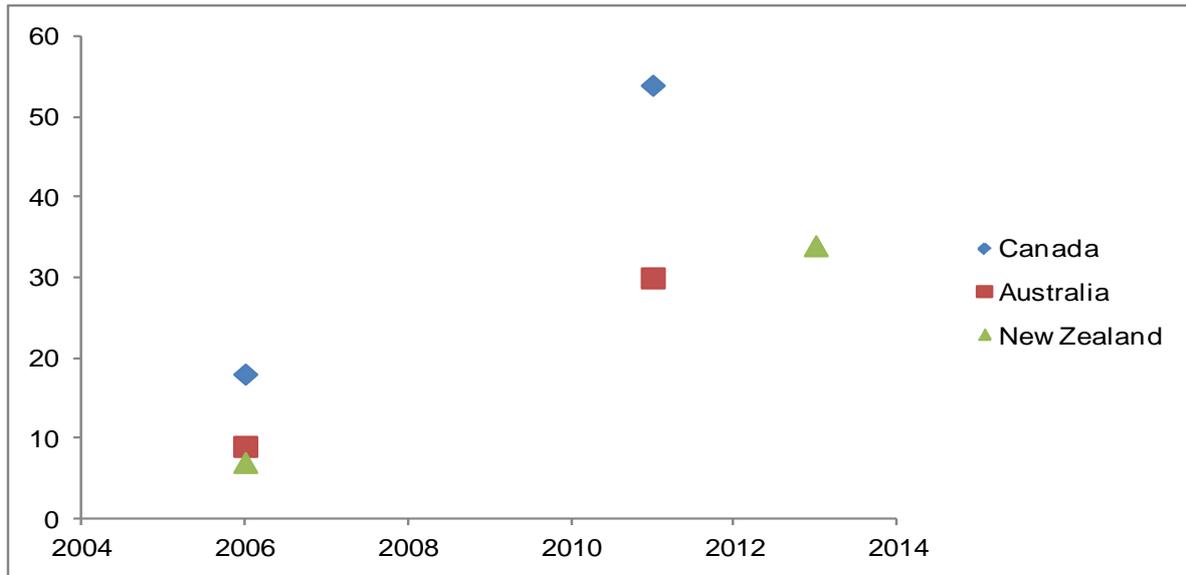
Next >>

Source American Statistical Association

Online Data Collection in the Census

Of the eight countries in Table 1 three have had a second Census which offered an online response mode. Figure 4 show that the three countries significantly increased the proportion of responses made online in the second Census in which an online option was available. In New Zealand the proportion of online responses increased from 7 per cent to 34 per cent in 2013^{xiv}, for Canada comparative figures are 18 per cent and 54 per cent and for Australia nine per cent and 31 per cent.

Figure 4 Proportion of Census Returns via the Internet Across Two Censuses.



Sources; Australian Census Bureau, Statistics Canada, Statistics New Zealand

Note: The 2011 New Zealand Census was delayed until 2013.

7.0 Benefits of an Internet Data Collection Mode

This section of the report examines the benefits that an internet data collection mode may bring. Benefits cover such areas as cost, data quality, respondent burden and data privacy and security. Census developments will not happen in isolation, indeed the nature of an online data collection mode makes it possible to develop a base model that can be modified for a specific survey other than the Census with relatively low marginal cost. The benefits outlined in this section are made between a paper based environment and the working assumptions on the delivery of Census 2021 which include an online response mode in Northern Ireland.

A face-to-face interview provides the most accurate survey responses. An interviewer is there to guide the respondent through the questionnaire ensuring answers are appropriate. However face-to-face is the most expensive data collection

Online Data Collection in the Census

mode. It is hoped that an internet mode will provide a level of accuracy tending towards that which a face-to-face interview would provide at a lower cost.

It would seem intuitive that responding to a survey by the internet would automatically reduce the cost. However there is no guarantee that increasing the proportion of Census returns made via the internet automatically leads to a reduction in cost. The initial set up, fixed costs, to develop an internet mode are significant. A critical number of returns is needed in order to realise a cost saving. Canada estimated that internet returns in 2011 would need to be in the region of 35-40 per cent to realise a cost saving. In fact in 2011 Canada achieved an internet returns rate of 54 per cent.

The position in Northern Ireland will be affected by whether NISRA runs a separate operation in Northern Ireland or, as in 2011, participates in a joint operation with ONS. If NISRA runs a separate operation, it is likely that a higher number of online returns would be required to cover the fixed costs associated with independently developing an online response mode than if NISRA, in partnership with ONS, develop an online response mode.

The need to maximise cost savings in an area where savings are not guaranteed would rationalise the option to provide an online response option for the wider NISRA survey programme. There is also a need to future proof, as far as possible, the online response mode in order that the platform could be used beyond 2021 in order to maximise cost savings.

Statistics Canada^{xv} suggested that cost savings arise in respect of:

- Reduced postage;
- Reduced mail – back and registration of paper questionnaires; and
- Reduced item non response, this was achieved in Canada by making core questions mandatory.

In addition to the benefits above Statistics Canada also realised benefits in relation to:

- High quality data;
- Low item non – response;

Online Data Collection in the Census

- Simplified processing of returns;
- Data in machine readable format; and
- Environmentally friendly in nature.

ONS^{xvi} expect cost savings in light of an internet mode to be associated with:

- Reduced calls to the contact centre for help
- Savings in paper scanning / capture; and
- Speedier capture of data.

As well as a cost saving the benefit relating to scanning also increases data quality. Online data capture removes scanning error which can be subject to issues such as illegible handwriting.

The US Census Bureau envisage, when the US short form Census goes online in 2020, cost savings in light of:

- Reduction in printing;
- Reduction in mailing; and
- Reduction in data capture.

The Minnesota Population Centre^{xvii} estimated that a 75 per cent internet response rate in traditional mail out / mail back areas of the US could realise cost savings due to reduced postage, printing and scanning in the region of \$2bn.

Further research in Northern Ireland is needed to inform the debate around the nature and magnitude of cost savings associated with an online response mode as part of a multi-mode Census design.

Editing comprises the systematic inspection of invalid and inconsistent survey responses and subsequent manual or automatic corrections. An online data capture mode makes the editing process less resource intensive than a paper questionnaire. An electronic questionnaire can be developed that only allows responses to a question that are within a specific range. If the questionnaire is electronic it is possible to “program” the questionnaire so that values outside the range are not accepted and the respondent cannot progress until a valid response is entered.

If we consider the above process as a “within” question edit it is also possible in an electronic questionnaire to incorporate a “cross” question electronic edit. For example if in one question the respondent indicates that she / he has one child but then enters “not applicable” when asked for the child’s age then an electronic questionnaire can recognise that the age of the child should have been a numeric value between 0 and 16 and flag an error in the response. The respondent can be prevented from progressing until the error is resolved.

Several Census questions require the respondent to “tick one box”. In other words select a response from a menu of multiple answers. With a paper questionnaire, self administered, it is possible that the respondent could mistakenly tick more than one box and continue the questionnaire. With an electronic questionnaire an error message can be programmed to appear on screen and the respondent cannot complete the questionnaire without rectifying the mistake. With an online data collection mode data entry is done directly by the respondent and edits applied in real time. The reduction in human interaction through direct data entry by the respondent should reduce cost and have a positive impact on data quality. Not all questions are applicable to all respondents in the Census or indeed in surveys in general. A paper questionnaire cannot “hide” the question(s) that are not applicable to the respondent. An electronic questionnaire can be programmed so that if a response to one question implies that other questions should not be answered then the respondent is *routed* to the applicable questions only. The ability to route the respondent will decrease the burden on the respondent by reducing the time taken to complete the Census or survey.

The 2011 Census showed large numbers of people in Northern Ireland whose first language was not English. The sole use of paper questionnaires would entail the respondent requesting a questionnaire in the language required. An electronic questionnaire can have the facility to move between languages and an online help function can be multi-lingual. Respondents in Wales had the ability to toggle between English and Welsh language questionnaires in 2011.

The success of an online data collection mode will depend to some extent on how concerns about data privacy and security are addressed. Such concerns are widespread; a survey on behalf of Big Brother Watch^{xviii} found that 79 per cent of

Online Data Collection in the Census

people globally are concerned about their online privacy. The public needs clear information on measures taken with regard to safeguarding their data and the benefits the internet brings. Privacy and security benefits occur due to a lesser degree of human interaction with an online form compared to a paper questionnaire. Some people may prefer completing their return online, viewing it as more secure than a paper return.

The exact method of contacting households for Census 2021 in Northern Ireland has not yet been decided, however the working assumption is that initial contact will be by mail and include an access code to complete the Census online. For any Census or survey that involves initial contact by mail an address register is a key asset. In Northern Ireland the most comprehensive address register, POINTER, was first used in Census 2011. With an online response mode and an electronic response/records management system (RMS) returned addresses can be electronically checked against POINTER in real time by enumerators in the field. Additionally enumerator workloads can be amended to reflect outcomes in the field in real time, therefore making more efficient use of the enumerators' time. Enumerators can be linked to the RMS and Census office by mobile devices.

The cost of adding an additional question and / or response category to an existing question is less in an online environment than in a paper questionnaire although it should be noted that the relationship between cost and questionnaire size is not straightforward.

The National Statistics Institute of Spain^{xix} published a SWOT (Strengths, Weaknesses, Opportunities and Threats) analysis that provides a useful summary of many of the characteristics of an internet response mode. Figure 5 shows the SWOT analysis.

Online Data Collection in the Census

Figure 5 SWOT Analysis for an Internet Response Mode.

Strengths (S)	Weaknesses (W)	
<ul style="list-style-type: none"> • Existence of only fixed overheads • Experience: it was possible to answer by Internet in 2001 Census and 2009 Pilot Census • Quality of information • Dissemination of data in a short period of time • Respondent burden reductions 	<ul style="list-style-type: none"> • Web application has to be safe, solid and with an attractive design • Integration among different channels • Security of the Web application • Simultaneousness of paper and Internet 	I N T E R N A L
<ul style="list-style-type: none"> • More and more people have an Internet connection • According to the economic global situation: minimizing costs • All the population will know a Census will be carried out in 2011 	<ul style="list-style-type: none"> • It is not possible to predict exactly the Web rate of response • People's knowledge of technology and familiarity with Web applications • Internet is not available in the 100 per cent of dwellings 	E X T E R N A L
Opportunities (O)	Threats (T)	

Source: National Statistics Institute, Spain

8.0 Suggested Design Artefacts for an Online Data Collection Mode

This section of the report outlines some design artefacts that have been successfully incorporated into various Censuses.

In 2006 Statistics Canada began to look for ways to promote the online response option. In advance of Census 2006 an Internet Response Promotion (IRP) study^{xx} was undertaken. The purpose of the study was to promote internet responses by encouraging households to complete their questionnaire online. To minimise the burden of data collection Statistics Canadian firstly devised a *model* the aim of which was to target households that were more likely than others to respond online. The targeting was based on households whose members used the internet. As there is no directory of internet users in Canada the source information for the model was the characteristics of those households that responded online to a previous Census Test in 2004. Using this information the model assigned likelihoods of households responding online. For the IRP study this resulted in a sample of 40,000 households being selected that were most likely to respond online. This sample was then split into two equal sub samples. One sub sample - the IRP sample - received a letter containing the URL of the Web site, a unique access code and a toll-free telephone

Online Data Collection in the Census

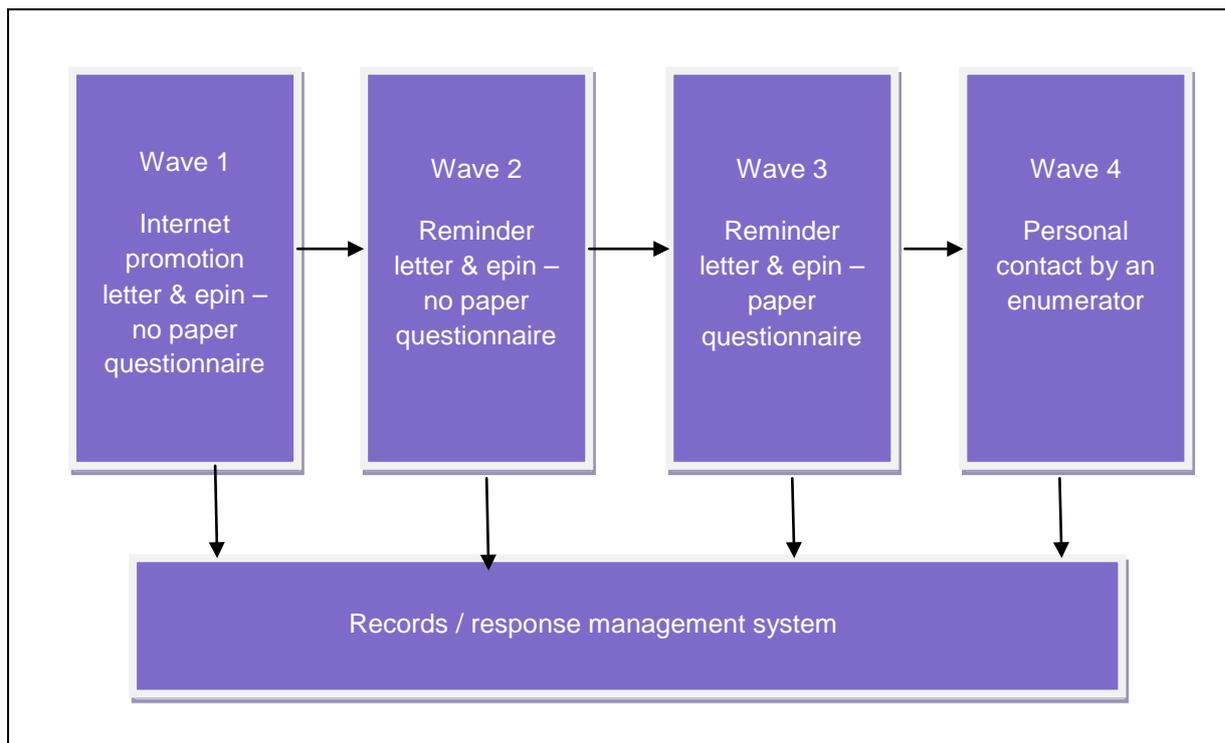
number. Respondents could call the number to get help or to ask for a paper questionnaire if they preferred. The other sub sample – the control group - received a paper questionnaire. The initial means of contact, either a letter only or a paper questionnaire, was referred to by Statistics Canada as the *method*. The method, as it comprised two means of initial contact, allowed Statistics Canada to quantify the impact of the means of initial contact on the level of internet responses.

The effect of the model, i.e. identifying households likely to respond via the Internet, can be measured by comparing the proportion of online responses in the Census with the proportion in the control group, since the households in the control group were all likely to respond online. The effect of the method - sending a letter only - can be measured by comparing the proportion of online responses for the IRP sample with the proportion for the control group, since only the former received a letter instead of a questionnaire. In summary the largest effect identified was that issuing just a letter with an internet access code increased the online response rate by 38 percentage points as against issuing a paper questionnaire. In comparison the targeting of households likely to respond online increased the online response rate by 5.5 percentage points.

The main element of the Canadian approach was a wave design as shown in Figure 6 below. The use of a letter only in waves 1 and 2 pushed respondents to use internet as their preferred mode of response. The pertinent feature of the Canadian wave design is the initial restriction of choice in the data collection mode.

The strategy in Canada as well as setting an overall response rate target aimed to ensure uniform response levels by region. It was put in place to make optimum use of the human and financial resources in the pursuit of response objectives. The strategy required the availability of management information on the progress of collection almost in real time. In other words an electronic RMS is a necessity.

Figure 6 The Canadian Wave Methodology.



Source; Author's graphical interpretation of Statistics Canada text^{xxi}. Note the RMS was added by the author.

Ever since the inception of the Portuguese 2011 Census an online response mode was envisaged. As referenced previously this response mode accounted for a surprising 50 per cent of Census returns. As in Canada the Portuguese employed a wave methodology with the focus being the online mode. Specific design artefacts included:

- Pre Census advertising campaign with a focus on the online response mode;
- Dedicated helpline with the technical capability to advise on the online mode;
- Week 1 online responses only;
- Week 2 online and paper with the preference for online;
- Week 3 paper based responses only;
- Strong commitment of the entire field force; and
- Remuneration independent of response mode.

The Spanish National Statistics Institute¹⁰ suggested the following ideas to increase online responses:

Online Data Collection in the Census

- Advertising campaign in “classic media” such as television, radio or newspaper;
- Advertisements in Web pages, outsourcing to experts in viral publicity^{xxii};
- Advertising in social networks;
- Careful composition of the letter sent to each dwelling; and
- Making it easier for people that do not have an Internet connection at home and are willing to answer the Census by internet to have access to an internet enabled computer.

The experience of the countries outlined above has provided some base design principles that may be utilised. The need for pre survey publicity of the online mode is a necessity. The public should also be made aware of the legislation that protects their data. The publicity may also inform the public that NISRA staff, involved in the handling of Census data, had security clearance and the nature of the physical infrastructure to ensure data security.

An electronic response / records management system (RMS) is recommended. An electronic RMS links the fieldforce, via mobile devices, to the Census office. Addresses can be checked, workloads distributed / amended in real time.

The previous design artefacts related to the “delivery” of the Census. Functional design artefacts, i.e. related to the questionnaire, were developed by the US Census Bureau^{xxiii} and are repeated below:

- Reproduces the survey questions and instructions in an electronic format;
- Allows response entry through data keying or data entry controls (radio buttons, check boxes, pick lists, etc.);
- Provides both sequential or non-sequential navigation;
- Incorporates auto-filled and auto-calculated fields;
- Performs branching and skipping of items based on response values;
- Uses edit tests to verify the response values for accuracy, integrity, completeness and shows error or warning messages;
- Uses field checks to test the validity of the response type (alphabetic, numeric, date, phone number, etc.) and shows error or warning messages;

Online Data Collection in the Census

- Provides final survey review facility to check for missing responses and unresolved edit failures;
- Allows respondents to print their response values;
- Offers a means for respondents to provide free-form comments;
- Provides context-sensitive help and general information about the survey;
- Allows the respondent to save the data already entered and to retrieve these data upon resuming the reporting session at some later time;
- Produces output that can be integrated with the survey processing system;
- Passes usability testing;
- Compatible with Web browsers; and
- Requires zero (or minimal) user configuration.

If a Government system were to experience operational difficulties in delivering an online response mode then citizens would lose faith in Government's ability to deliver large scale projects. In order to mitigate against the possibility of the online platform crashing under the weight of demand it is recommended that online responses are staggered by means of issuing the invitations to respond across several time periods. This staggering could coincide with the wave methodology aimed at increasing online responses.

To make a success of an internet data collection mode requires steps to be taken prior to Census data capture (e.g. Spain and Portugal) and also in the actual design of the internet application (US Census Bureau).

9.0 Conclusions

Census developers and social researchers are at a critical juncture in determining data collection modes. Internet data collection is technically feasible, but the initial investment in hardware and software is costly. Given the divide in computer knowledge and access, internet data collection is viable for some, but not for all. Therefore the internet cannot fully replace the existing paper questionnaire, at this point in time.

The design proposed for Census 2021 in relation to response options is multi-mode with an emphasis on an online mode.

Online Data Collection in the Census

In order to reduce the impact of the digital divide Government's digital strategies should be beneficial to an online Census / survey data collection mode. Strategic objectives aim to increase electronic literacy, remove cost barriers and set minimum delivery standards.

The acceptance of new technologies in statistical investigations is developing at a significant pace. The statistical arms of the UN and the EU as well as many national statistical institutes promote the use of new technologies including the use of an online data collection mode for the Census and surveys. To not offer an online response mode in the Census and surveys would therefore be out of step with this direction of change.

While the level of internet access in Northern Ireland is high, the level is below EU and UK averages. This implies that to have a predominantly online Census in 2021 will be more difficult in Northern Ireland than in the other countries of the UK.

The potential benefits from alternative modes or mixed modes of data collection are seen to include: reduced cost (cheaper than face-to-face interviewing); increased coverage; reduced response and selection biases compared with single mode surveys (when not using a comprehensive sampling frame); increased response rates; increased sample representativeness; reduced measurement error; and improved timeliness of an online mode.

In Northern Ireland the online mode offered in Census 2011 was a good foundation for further development. As shown in section 7, publicity prior to any fieldwork highlighting the benefits of an online mode is recommended. It would be fair to say that NISRA did not publicise the online mode to any great degree prior to Census 2011. A publicity campaign, which as well as introducing the online mode, covers aspects relating to data privacy and security is essential. Other positive aspects of an online mode may also be portrayed for example the environmentally friendly nature of online i.e. reducing the use of paper.

A wave design such as that employed in Canada is recommended. The key design objective should be to restrict respondent choice and push the respondent to respond online. Also the need to stagger responses in order to mitigate against a potential system crash as a result of the magnitude of demand is required.

Online Data Collection in the Census

Remuneration of enumerators should be carefully considered in light of the outcome in Portugal. However a balance is needed between overall enumeration and the objective to minimise costs.

While the focus of the report has been an online mode via the web it is apparent that restricting an online mode to the web alone would miss a significant opportunity to avail of modern digital technology. Already smartphones and tablets have been taken up by early adopters and their market share will only increase. Therefore it is necessary that the development of a digital data collection mode for the Census and the wider NISRA survey programme should include emerging technologies and keep a watching brief for yet unknown digital products. There is also a need to future proof, as far as possible, the online response mode in order that the platform could be used beyond 2021 in order to maximise cost savings.

Restricting the development of an online mode to the Census would represent a missed opportunity to avail of economies of scale. While the Census is the primary focus of this paper it is recommended that NISRA's wider survey programme be considered in respect of the utility of providing an online response mode.

In common with the Skinner review more research is recommended. Specifically NISRA, before offering an online mode, should undertake further research into:

- Mode effects;
- Cost; and
- Public attitudes to an online mode in the Census and the wider NISRA survey programme.

This paper has provided a brief introduction to the possibility of an online data collection mode for the Census. The objective of the paper is to gain a consensus to offer a digital responses mode in the Census. While the report has introduced some design suggestions it by no means provides a definitive guide to designing an online response mode and should be considered as a starting point for further more detailed research into an online data collection mode.

Notes and References

- ⁱ Government Digital Strategy <https://www.gov.uk/government/publications/government-digital-strategy>
- ⁱⁱ Digital Inclusion Strategy <https://www.gov.uk/government/publications/government-digital-inclusion-strategy>
- ⁱⁱⁱ For more information visit the website <http://www.go-on.co.uk/get-involved/ni/>
- ^{iv} Digital Transformation Service – Citizen Contact Strategy <http://www.dfpni.gov.uk/digital-transformation-service-citizen-contact-strategy>
- ^v National Statistician's Recommendation <http://www.ons.gov.uk/ons/about-ons/who-ons-are/programmes-and-projects/beyond-2011/beyond-2011-report-on-autumn-2013-consultation--and-recommendations/index.html>
- ^{vi} Chris Skinner, John Hollis and Mike Murphy (2013) "*Beyond 2011: independent review of methodology*" accessed at <http://www.ons.gov.uk/ons/about-ons/who-ons-are/programmes-and-projects/beyond-2011/reports-and-publications/index.html>
- ^{vii} <http://www.nrscotland.gov.uk/news/2014/plans-for-a-census-in-scotland-2021>
- ^{viii} National Statistics Board Minutes of Meeting of the 20th May 2014. Accessed at <http://www.nsb.ie/publications/minutesofnationalstatisticsboardmeetings/nsbminutes20may2014/>
- ^{ix} OFCOM (2014) "*Adult's Media Use and Attitudes 2014*" accessed at http://stakeholders.ofcom.org.uk/binaries/research/media-literacy/adults-2014/2014_Adults_report.pdf
- ^x Respondents could give more than one reason for not accessing the internet and therefore totals will not sum to 100 per cent
- ^{xi} United Nations Statistics Division (2011). "*Report on the Results of a Survey on Census Methods used by Countries in the 2010 Census Round, Working Paper UNSD/DSSB/1*" accessed at <http://unstats.un.org/unsd/censuskb20/KnowledgebaseArticle10696.aspx>
- ^{xii} United Nations Economic Commission for Europe (2012), "*Conference of European Statisticians, Group of Experts on Population and Housing Censuses, 2011 Census in Portugal – implementation of the e-Census system, Statistics Portugal*" accessed at http://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.41/2012/census_meeting/8_Portugal_Eng.pdf
- ^{xiii} Castro D "*e-Census Unplugged: Why Americans Should Be Able to Complete the Census Online*" The Information Technology & Innovation Foundation, (2008), Washington DC. Accessed at <http://www.itif.org/files/eCensusUnplugged.pdf>
- ^{xiv} The Census in New Zealand was due to have taken place in 2011. However due to an earthquake in Christchurch the Census was postponed until 2013.
- ^{xv} Statistics Canada, (2012), "*Overview of the Census, Census Year 2011*" accessed at <http://www12.statcan.gc.ca/census-recensement/2011/ref/overview-aperçu/pop1-eng.cfm>
- ^{xvi} Office for National Statistics, (2012) *Census 2011 Update*. Accessed at <http://www.ons.gov.uk/ons/guide-method/census/2011/index.html>
- ^{xvii} Ruggles S "*Comment on Census Internet and Electronic Data Collection*" Minnesota Population Centre (year not cited)

Accessed at https://www.census.gov/cac/about_census_advisory_committees/pdf/ruggles_comment-electronic_data_collection.pdf

^{xviii} Report accessed at <http://www.bigbrotherwatch.org.uk/home/2013/06/new-research-global-attitudes-to-privacy-online.html>

^{xix} United Nations Economic and Social Council Conference of European Statisticians Group of Experts on Population and Housing Censuses (2009) "*Online Data Collection Theoretical and Practical Analysis about Census by Internet A Note by the National Statistics Institute, Spain*". Accessed at <http://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.41/2009/18.e.pdf>

^{xx} Anne-Marie Cote, Dannielle Laroche (2009) "*Statistics Canada International Symposium Series: Proceedings; The Internet: A New Collection Method for the Census*" Statistics Canada 2009. Accessed at <http://www.statcan.gc.ca/pub/11-522-x/2008000/article/10986-eng.pdf>

^{xxi} United Nations Economic and Social Council Conference of European Statisticians Group of Experts on Population and Housing Censuses (2012) "*Internet data collection in the Canadian census of population*", Statistics Canada, 2012. Accessed at http://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.41/2012/census_meeting/6_Canada_Eng.pdf

^{xxii} "Viral Publicity" can be defined as a strategy to attract attention to one's company, whilst causing fans to spread one's message by themselves, thereby creating new uses of one's product, new slangy slogans, new communities of shared interest and new ways of being who one thinks.

^{xxiii} Howard Kanarek and Barbara Sedivi (1999) "*Internet Data Collection at the US. Census Bureau*". U. S. Census Bureau 1999. Accessed at <http://www.fcsn.gov/99papers/kanarek.pdf>