

Unlocking the Past: the 1901 census online

2nd Parish of <i>St. Andrew's</i>		Parish of <i>St. Andrew's</i>		3rd Parish of <i>St. Andrew's</i>		4th Parish of <i>St. Andrew's</i>		5th Parish of <i>St. Andrew's</i>		6th Parish of <i>St. Andrew's</i>		7th Parish of <i>St. Andrew's</i>		8th Parish of <i>St. Andrew's</i>	
Household No.	Street	Name and Surname of Head of Family	Relation to Head of Family	Sex	Age	Profession or Occupation	Employed	Employed	Wages	Wages	Wages	Wages	Wages	Wages	Wages
1	<i>St. Andrew's</i>	<i>John Christie</i>	<i>Head</i>	<i>Male</i>	<i>45</i>	<i>Assistant</i>			<i>10s</i>	<i>10s</i>	<i>10s</i>	<i>10s</i>	<i>10s</i>	<i>10s</i>	<i>10s</i>
		<i>Margaret Christie</i>	<i>Wife</i>	<i>Female</i>	<i>43</i>	<i>Domestic</i>	<i>X</i>		<i>5s</i>	<i>5s</i>	<i>5s</i>	<i>5s</i>	<i>5s</i>	<i>5s</i>	<i>5s</i>
		<i>James Christie</i>	<i>Son</i>	<i>Male</i>	<i>15</i>	<i>Domestic</i>			<i>5s</i>	<i>5s</i>	<i>5s</i>	<i>5s</i>	<i>5s</i>	<i>5s</i>	<i>5s</i>
		<i>John Christie</i>	<i>Son</i>	<i>Male</i>	<i>12</i>	<i>Domestic</i>			<i>5s</i>	<i>5s</i>	<i>5s</i>	<i>5s</i>	<i>5s</i>	<i>5s</i>	<i>5s</i>

REPORT BY THE COMPTROLLER AND AUDITOR GENERAL
HC 1259 Session 2002-2003: 14 November 2003

executive summary

Introduction and main findings

- 1 The Public Record Office is a government department in its own right established under the Public Records Acts of 1958 and 1967, and the Chief Executive, the Keeper of Public Records, reports directly to the Lord Chancellor. It became an Executive Agency on 1 April 1992. The Public Record Office, located at Kew, administers the public record system of the United Kingdom, and advises government on public record issues and related information matters. It was also jointly responsible with the Office for National Statistics for the Family Records Centre (located in Islington) which houses microform records of decennial censuses of population from 1841 to 1901, and the indexes of births, deaths and marriages in England and Wales from 1837. In April 2003, the Public Record Office became the National Archives for England, Wales and the United Kingdom¹, combining the roles of the Public Record Office and the Historical Manuscripts Commission.



- 2 The 1901 census website - www.census.pro.gov.uk - was officially launched by the Public Record Office at 9 am on 2 January 2002. The service was designed to provide access to 1 million users, with a peak of 1.2 million users, in a 24 hour period. However, by midday on 2 January 2002, 1.2 million users per hour were attempting to access the site from locations across the world. Between 2 and 6 January 2002, the site continued to experience 1.2 million users per hour, overwhelming the site. On 7 January 2002, the Public Record Office and its contractor, QinetiQ², agreed to close the site to general Internet access to allow QinetiQ to undertake a technical investigation. The website was released to the public on a limited basis on 6 August 2002 and was made fully available to the public on 21 November 2002, since when it has operated effectively.

¹ As this report covers the period prior to the creation of the National Archives, we refer throughout the report to the Public Record Office.

² At the time the contract was originally let, the Public Record Office was dealing with the Defence Evaluation and Research Agency (DERA). In July 2001, a major part of the business of DERA was transferred to a new group of companies called QinetiQ plc. The census contract was transferred to a company within the group called QinetiQ Ltd, a wholly owned subsidiary of QinetiQ plc. The Ministry of Defence retains a 67½ per cent shareholding in QinetiQ plc. Since 28 February 2003 QinetiQ is no longer part of the public sector. Throughout this report "QinetiQ" is used to refer to either QinetiQ Ltd or DERA on the basis that there was operational continuity between the two entities as far as dealings with the Public Record Office were concerned.

3 This report examines the basis for the project, the problems encountered during its implementation, and the steps taken to restore services to the public. It also looks at lessons learned with general relevance for the delivery of central government services over the Internet. Our main findings are:

- The 1901 census online is now a successful project which receives between 8,000 and 10,000 visitors per day and which has generated gross revenues of £4.5 million in the the period up to 31 October 2003. Representatives from the world of genealogy contacted by the National Audit Office were broadly positive about the service the website now provided.
- The Public Record Office managed the main risks associated with the project - for example, by recognising that it did not have all the skills in-house to undertake the project alone, by transferring development risk to a contractor, and by completing research into potential levels of demand. The Public Record Office could not, however, transfer the risks to its reputation.
- While the Public Record Office developed a pre-launch strategy based on a low key launch, the actual level of press interest generated by the launch was unexpectedly high. And the site was launched on 2 January 2002, in a holiday period, increasing the number of people who could access the site from home. These two events were contributory factors to the creation of the high level of demand that exceeded the site's capacity.
- The Public Record Office's experience in making the 1901 census available online has lessons for other government departments and agencies considering a similar method of service delivery. These lessons are summarised at the end of the Executive Summary on page 7.

Appendix 1 to this report outlines our approach to the study; Appendix 2 shows the chronology of key events; and Appendix 3 provides a glossary of terms.

The Public Record Office decided to make the 1901 census data available on the Internet

4 The Public Record Office had a statutory duty to make census information available to the public following the 100 year period during which records were closed to public access. In the past, census information has been made available on microfiche and microfilm at the Family Records Centre. To improve public access, widen social inclusion, cater to a growing worldwide interest in genealogy and contribute to the modernising government agenda, the Public Record Office decided to make the 1901 census data available on the Internet, in addition to making it available in microfiche format.

The service would be developed and delivered by a contractor

5 The Public Record Office did not have the investment funds available to achieve the above objective. It therefore decided to tender this work on the basis that a contractor would fund the development costs and recoup them from revenue generated from access to the website. Putting the 1901 census online was a three stage process involving: the digitisation of the census data; the design of the website; and the operation and ongoing maintenance of the site, including a contact centre to provide advice for users.



Responsibility for the project was divided between the Public Record Office and QinetiQ

- 6 The Public Record Office let a ten year contract with QinetiQ, following a tendering process conducted in accordance with guidelines issued by the Central Computer and Telecommunications Agency³ and the Office of Government Commerce. Responsibility for the project was divided between the Public Record Office and QinetiQ. The Public Record Office, with the permission of the Registrar General of the Office of National Statistics, made the 1901 census microfilms available to QinetiQ at agreed times to allow digitisation of the images. QinetiQ's role was to design, develop, implement, commission, operate and maintain an online access service which would also collect the revenues due from the use of the service. The Public Record Office retained ownership of the census data and associated intellectual property rights. QinetiQ retained ownership of the hardware and software used for the provision of the service. Under the arrangements, QinetiQ provided all the development investment and bears the operational costs of the service. The Public Record Office incurred only contract management and administrative costs.
- 7 Access to the census website and initial search facilities are provided free. More detailed review of data and images of the census returns are subject to a charge. QinetiQ retains revenues at an agreed level to cover its operating costs and all further annual net revenues until it has fully recovered its agreed development costs. Any additional net revenues are then divided equally between QinetiQ and the Public Record Office.

Problems arose during the project's implementation

- 8 During 2000, the Public Record Office and QinetiQ identified that the timetable for capture of the census data was under threat compromising the statutory opening day of 2 January 2002. Following detailed discussions and with the agreement of QinetiQ and its subcontractor, Enterprise and Supply Services (a division of HM Prison Service), the Public Record Office took over the operational management of the data capture process to enable data to be available in sufficient time for the website launch on the target date. QinetiQ and Enterprise and Supply Services retained contractual control. This decision involved the appointment of new subcontractors to carry out the data capture work and required the Office of National Statistics' approval for the data to be processed overseas. The Public Record Office was responsible for quality control of the work; QinetiQ was responsible for functionality and data integrity checks. The appointment of subcontractors to complete the transcription resulted in a net cost to HM Prison Service of £1.8 million.
- 9 QinetiQ retained complete responsibility for the technical development of the website and the Public Record Office agreed that the launch should go ahead on the basis of assurances from QinetiQ that the technical development was complete and could cope with the agreed levels of expected loading of the system. The website was made available to the public on 2 January 2002. While the Public Record Office had planned a low key launch for the site, this did not materialise due to the levels of press coverage generated in the days immediately after the launch. The site was withdrawn on 7 January 2002 due to pressure from a surge in initial demand from users, which amounted to 1.2 million users an hour compared to the forecast maximum load of 1.2 million users a day. The site was not able to divert this level of excess demand.


³ From 1 April 2001, the Central Computer and Telecommunications Agency became part of the Office of Government Commerce.

- 10 The system could not be restored immediately to general Internet access. However, it continued to be accessible at some local record offices and libraries, and at the Family Records Centre and the Public Record Office until 6 February 2002. Following a review by QinetiQ and consultants appointed by the Public Record Office (with the assistance of the Office of the e-Envoy), QinetiQ completed some further work on the site, including enhanced firewalls and enhanced load balancer configurations, which were designed to divert excess demand. Subsequent to QinetiQ testing of this further work, the Public Record Office instigated its own programme of testing, which was carried out by an external contractor, Keynote.
- 11 The Public Record Office and QinetiQ's calculation of planned capacity of 1.2 million users over a 24 hour period was based on independent market research into other genealogical sites offering similar services, and the results of their pilot project focused on 1891 census returns for the County of Norfolk. While the market research provided some indications about comparable sites, direct comparison was very difficult due to variations in the number and the nature of the specific records made available. However, the FamilySearch website run for the Genealogical Society of Utah launched in May 1999 with 640 million personal records received four times the expected demand and crashed at launch. After relaunch, that site was receiving usage well within the site capacity. The experience with the 1901 census website has been similar since its relaunch in November 2002.
- 12 Between April and August 2002, the Public Record Office and QinetiQ (facilitated by the Office of the e-Envoy) attempted to reach agreement about the robustness of the system. By 27 August 2002 both parties agreed that the site could be made available to the public on a restricted basis. The formal opening hours were progressively extended, although in practice the site was often available 24 hours a day. After residual problems were resolved, the site became fully operational with 24 hour access on 21 November 2002.

The project was designed to be self-financing

- 13 Under the contract, QinetiQ's agreed development costs were £8.15 million, to be recovered by QinetiQ from revenues before any income is received by the Public Record Office. QinetiQ spent further sums during 2002 and has lodged a claim with the Public Record Office for a corresponding increase in the agreed development and operational costs, on the grounds that this additional expenditure was for enhancements to the website. The Public Record Office maintains that the additional work was required to fulfil the specification in the contract. The Public Record Office and QinetiQ are currently in negotiations on this issue.



- 
- 14** The Public Record Office had incurred total net costs of £1.7 million. After recovery of its agreed development costs of £8.15 million, QinetiQ shares further net revenues equally with the Public Record Office. The site has generated total revenues of £4.5 million in the period up to 31 October 2003. QinetiQ told the National Audit Office that no clear trend had yet emerged in the monthly profile of receipts.

The experience in putting the English and Welsh 1901 census online contrasts with the project which put the Scottish 1901 census online

- 15** The circumstances surrounding the projects to place the English and Welsh 1901 census and the Scottish 1901 census online were very different. In particular, the scale of the Scottish project was much smaller and less complex than the project for England and Wales. Additionally, two key factors that might have controlled demand for online access to the Scottish 1901 census records were the upfront charge for access to the Scottish website and the progressive release of the data on microfilm before the release of the data online. Both factors reflect different policy decisions taken in England and Scotland.
- 16** Our examination shows that the Public Record Office, together with QinetiQ, were ultimately successful in this ambitious project after 11 months of difficulties. The initial crash was triggered by overwhelming demand that caused technical problems for a system which was designed for much lower capacity and which was not able to divert such a large initial surge of traffic. The resolution of these problems became protracted, due to the parties failing easily to agree on the results from their separate testing programmes, and the relationship between the parties deteriorated. The Office of the e-Envoy assisted to facilitate agreement between the parties over the basis on which the site could be made available to the public. The National Archives, formerly the Public Record Office, and QinetiQ are now working together effectively, and the site is generating a steady flow of revenue.
- 17** The National Archives should apply the lessons learned from this experience in its future dealings with third parties involving public access to data.

We recommend that the National Archives should:



NAME and address of each Person.	RELATION to head of Family.	AGE	SEX	COLOUR	RELIGION	EDUCATION	PROFESSION, OCCUPATION, TRADE, BUSINESS, or Vocation.	INDUSTRY or BUSINESS.	REMARKS.
Henry Jones	Head	45	M	White	Anglican	None	None	None	
John Jones	Wife	42	F	White	Anglican	None	None	None	
William Jones	Son	15	M	White	Anglican	None	None	None	
Elizabeth Jones	Daughter	12	F	White	Anglican	None	None	None	
Thomas Jones	Son	10	M	White	Anglican	None	None	None	
James Jones	Son	8	M	White	Anglican	None	None	None	
John Jones	Son	6	M	White	Anglican	None	None	None	
William Jones	Son	4	M	White	Anglican	None	None	None	
Elizabeth Jones	Daughter	3	F	White	Anglican	None	None	None	
John Jones	Son	2	M	White	Anglican	None	None	None	
William Jones	Son	1	M	White	Anglican	None	None	None	



- 1 Monitor closely the ongoing financial performance of the 1901 census website and, in particular, obtain independent assurance on the extent of operating costs incurred and revenues received by QinetiQ.
- 2 Work with QinetiQ to maximise the marketing opportunities and consequent revenue from the 1901 census website, for example, in promoting take up by users overseas.
- 3 As early as practicable, develop a forecast for the revenue profile from the site and determine a strategy for the deployment of future funds due to it.
- 4 Include in its corporate plan a clear strategy for its ongoing objectives for the site, including an early appreciation of how the completion of the current 10 year contract with QinetiQ is to be managed so as to maintain continuing financial benefit and public access.
- 5 Check that QinetiQ continues to respond to user reaction to the site, in particular in the areas of search capability and error correction.
- 6 Develop an early strategic consideration of its approach to the release of the 1911 census data, including:
 - to what other media the original records are to be transferred and how this is to be achieved;
 - how the reputational risk associated with the release of new data is to be managed;
 - how the potential consequences of the increased volume of images of the 1911 records are to be assessed;
 - the extent to which the necessary confidentiality requirements pre-launch can be accommodated within licensing arrangements currently operating, or whether a more specific contract such as for the 1901 census remains relevant;
 - in the light of the revenue earnings from the 1901 census, what mechanism for the launch of the 1911 census represents the optimum revenue earning opportunity for the National Archives, whilst achieving other objectives for access and social inclusion.

Lessons with wider application

The 1901 census project raises issues of relevance for managing central government services on the Internet.

Examples of best practice demonstrated by the project

- 1 Transfer the commercial development risk to the contractor, to be funded, where possible, from subsequent revenues earned by the service.
- 2 Retain the ability for ongoing commercial exploitation of the core data upon which the service is based, and maximise cost recovery by securing an interest in future revenues from the project.
- 3 Make a realistic assessment of data conversion requirements where these are essential to the success of ultimate service delivery.
- 4 Maximise the opportunity to improve the service being put online from first principles and develop a clear marketing strategy for the service to add value for the users and widen social inclusion through improved access to the service.



Lessons for wider application for managing services using the Internet

- 5 Recognise the distinction between commercial risk which may be transferred to third parties and reputational risk that may remain with the contracting party.
- 6 Maximise the clarity of contractor obligations so that both parties are clear as to the intended completion state.
- 7 Include in contracts the facility to obtain independent assurance on the rigour of pre-launch system testing by the contractor.
- 8 Before the service goes live, develop an agreed post-launch disaster recovery strategy.
- 9 Make a realistic assessment, as far as possible, of usage, and ensure that this assessment is supported by a capacity management strategy that will successfully divert unexpected overloads experienced in practice.
- 10 Where appropriate, carefully select both the rate and timing of launch of new services so as to maximise the opportunity to resolve unforeseen problems before peak demand has built up.