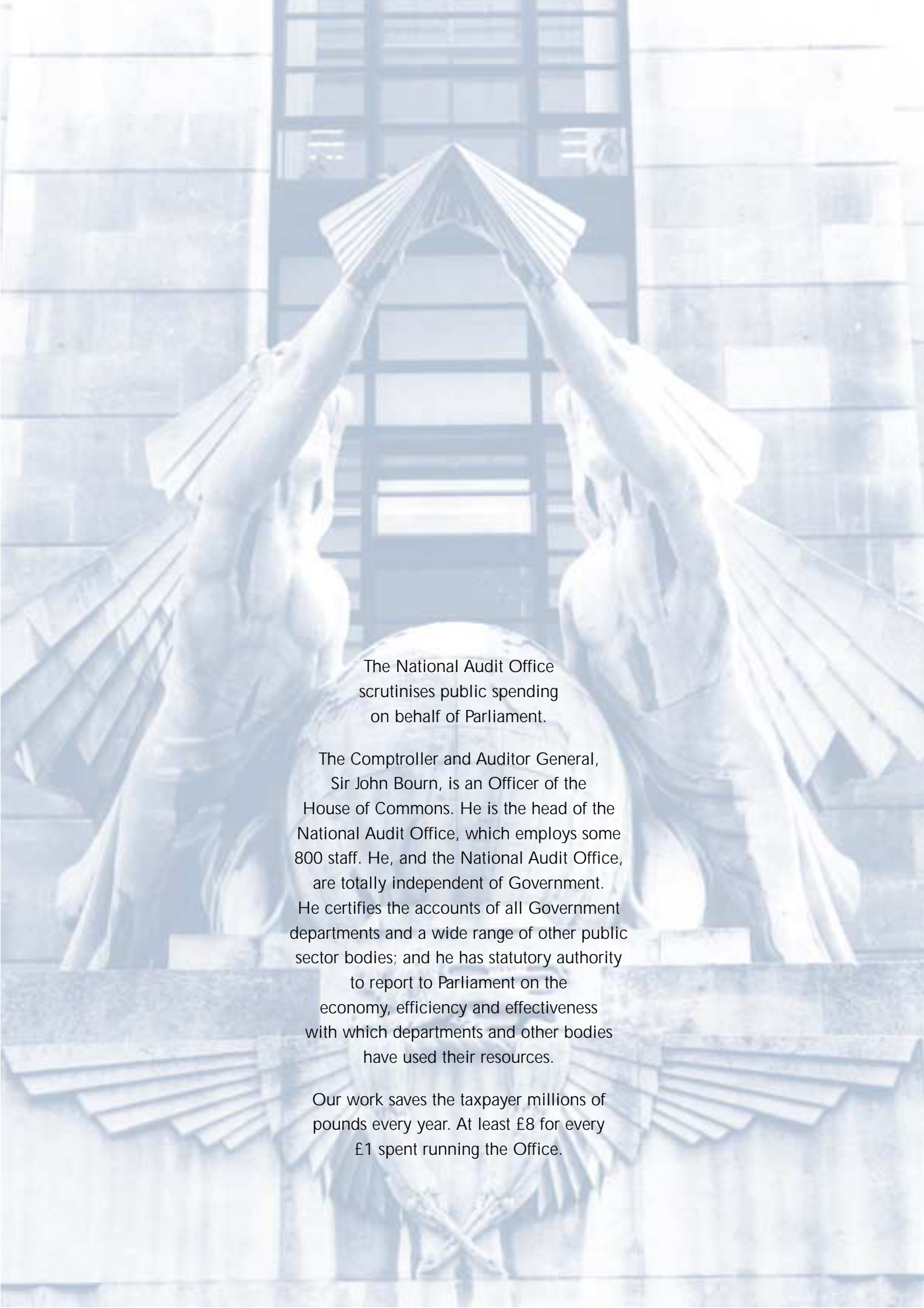


Unlocking the Past: the 1901 census online

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





The National Audit Office
scrutinises public spending
on behalf of Parliament.

The Comptroller and Auditor General,
Sir John Bourn, is an Officer of the
House of Commons. He is the head of the
National Audit Office, which employs some
800 staff. He, and the National Audit Office,
are totally independent of Government.
He certifies the accounts of all Government
departments and a wide range of other public
sector bodies; and he has statutory authority
to report to Parliament on the
economy, efficiency and effectiveness
with which departments and other bodies
have used their resources.

Our work saves the taxpayer millions of
pounds every year. At least £8 for every
£1 spent running the Office.

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>M</i>	<i>45</i>	<i>Assistant</i>			<i>10</i>	<i>10</i>	<i>10</i>	<i>10</i>	<i>10</i>	<i>10</i>	<i>10</i>
		<i>Margaret Christie</i>	<i>Wife</i>	<i>F</i>	<i>43</i>	<i>Domestic</i>	<i>X</i>		<i>10</i>	<i>10</i>	<i>10</i>	<i>10</i>	<i>10</i>	<i>10</i>	<i>10</i>
		<i>James Christie</i>	<i>Son</i>	<i>M</i>	<i>15</i>	<i>Scholar</i>			<i>10</i>	<i>10</i>	<i>10</i>	<i>10</i>	<i>10</i>	<i>10</i>	<i>10</i>
		<i>John Christie</i>	<i>Son</i>	<i>M</i>	<i>12</i>	<i>Scholar</i>			<i>10</i>	<i>10</i>	<i>10</i>	<i>10</i>	<i>10</i>	<i>10</i>	<i>10</i>
		<i>Margaret Christie</i>	<i>Daughter</i>	<i>F</i>	<i>10</i>	<i>Scholar</i>			<i>10</i>	<i>10</i>	<i>10</i>	<i>10</i>	<i>10</i>	<i>10</i>	<i>10</i>

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

This report has been prepared under Section 6 of the National Audit Act 1983 for presentation to the House of Commons in accordance with Section 9 of the Act.

John Bourn National Audit Office
Comptroller and Auditor General 10 November 2003

The National Audit Office study team consisted of:

Steve Allen, Gary Chan and Helen Thompson under the direction of Aileen Murphie

This report can be found on the National Audit Office web site at www.nao.gov.uk

For further information about the National Audit Office please contact:

National Audit Office
Press Office
157-197 Buckingham Palace Road
Victoria
London
SW1W 9SP

Tel: 020 7798 7400

Email: enquiries@nao.gsi.gov.uk



Contents

Executive summary 1

Part 1

The objective of going online	9
There is a statutory duty to make the 1901 census available to the public	9
The Internet was chosen to meet anticipated high demand for access to the 1901 census data	9
The launch date of 2 January 2002 was chosen to coincide with statutory requirements in order to meet user and government expectations	9
The work to set up the website was contracted out	9
QinetiQ was chosen through a publicly tendered competition	10
The project was planned to be self-financing	11
A cutting edge service was planned to be delivered	12

Part 2

Problems arising during implementation	15
Initially, insufficient progress was made converting census records into a digital format	15
Additional resources were required to complete the digitisation process	15
The Public Record Office stepped in to ensure quality control of data preparation	16
The website design process was QinetiQ's responsibility	17
A progress review concluded that the project should go ahead with its agreed launch date	18
The Joint Programme Board received regular reports on the project's progress	18
The Public Record Office and QinetiQ agreed a joint marketing strategy for the census website	18
When the 1901 census website was launched it became overwhelmed by a high level of demand from users	19

Part 3

How the implementation problems were resolved	21
After the crash, both the Public Record Office and QinetiQ instigated technical reviews	21
The Public Record Office and QinetiQ reached agreement on system changes with the assistance of the Office of the e-Envoy	22
The website became fully operational on 21 November 2002	23
The Public Record Office and QinetiQ shared the costs of putting things right	23
The relaunched website has been well received	24
Making the 1901 census available online has implications for future projects	25
Scottish 1901 census data online	26

Appendices

1. Our approach to the study	28
2. Chronology	29
3. A glossary of terms	32
4. Putting 1901 census data for Scotland online	34

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
Henry Jones	Head	45	M	White
Elizabeth Jones	Wife	40	F	White
John Jones	Son	15	M	White
Mary Jones	Daughter	12	F	White
James Jones	Son	10	M	White
William Jones	Son	8	M	White
Elizabeth Jones	Daughter	6	F	White
John Jones	Son	4	M	White
Mary Jones	Daughter	3	F	White
James Jones	Son	2	M	White
William Jones	Son	1	M	White



- 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.

John Bull's tailor measures him for a new suit - a satirical comment on a dramatic rise in population, as demonstrated by the 1901 census

Source: cartoon by F. Matania, 1901



Part 1

The objective of going online

1.1 This part of the Report sets out the Public Record Office's statutory duty to make the 1901 census returns available to the public. It explains how the decision to put the records on the Internet was taken, the way the project was set up and financed, and what service was planned to be delivered.

There is a statutory duty to make the 1901 census available to the public

1.2 The period during which the 1901 census returns remained closed to the public is governed by the Public Records Act 1958, section 5(1) and was extended through the Lord Chancellor's Statutory Instrument No 12 of June 1966 to 100 years. The records were therefore due to be made available to the public on 2 January 2002. The legislation does not set out how records are to be made available. The statutory requirement was met by making the census returns available free of charge on microfiche at the Public Record Office at Kew.

The Internet was chosen to meet anticipated high demand for access to the 1901 census data

1.3 On the basis of its experience when the 1891 census returns were opened to public access in January 1992, the Public Record Office anticipated a very large demand for access to the 1901 census returns, which would have caused lengthy queues and overcrowding at the Family Records Centre. In part to cope with this demand, but primarily to further the modernising government agenda, in particular to make public information available electronically by 2005, the Public Record Office decided in August 1998 to make the 1901 census returns available on the Internet, as well as at its offices.

1.4 The Public Record Office also considered that the timing was right to go online because of the massive growth in Internet usage since 1992. It saw an online service as widening social inclusion, because data would be accessible from home and in local libraries without the need for a physical visit to the Family Records Centre in Islington or the Public Record Office at Kew, both in London. The Public Record Office considered 1901 census data as the priority for digitisation, rather than earlier years which were already publicly available on microfiche, due to the pressure of additional demand for the new data.

The launch date of 2 January 2002 was chosen to coincide with statutory requirements in order to meet user and government expectations

1.5 The Public Record Office fulfilled its statutory requirements by making the 1901 census data available in microfiche format on 2 January 2002. The Public Record Office considered that it was essential to launch the website on the same date to meet user expectations. It also decided not to release the census gradually - for example, on a county-by-county basis - as this would have reduced the usefulness of the search facilities available, potentially provoked criticism of preference for one county over another, and would not have met user expectations.

The work to set up the website was contracted out

1.6 The Public Record Office had previous experience of managing information technology projects, through the development and operation of its computerised catalogue and its document ordering system. The Public Record Office's internal information technology function had experience of outsourcing the National Data Archive for Datasets to the University of London, and was able to

specify the conditions of an outsourcing contract. The Public Record Office also had experience in acquiring external funding through, for example, the Heritage Lottery Fund grant for microfilming World War I soldiers' records. However, it did not have direct experience of a Private Finance Initiative (PFI) contract.

- 1.7 The Public Record Office did not have the financial resources to fund the development of an online service. The Public Record Office considered using family history volunteers for the transcription elements of the project but concluded, regretfully, that this option was not viable given the nature and timing of the project and the need to maintain confidentiality. However, it decided, in June 1998, to consider the implications of making the 1901 census digitally available. It concluded that the project could be managed using PFI arrangements. To demonstrate further the benefits of such an arrangement compared to other options, a Public Sector Comparator exercise was commissioned from consultants (the DMW Group) in May 1999 to confirm, or otherwise, the Public Record Office's view of the project. The consultants considered seven options and concluded that a PFI solution should be adopted because it provided the opportunity for a full electronic service with effect from 2 January 2002, met the strategic aim of delivering census data in electronic format, and did not require significant investment by the Public Record Office.
- 1.8 The other six options considered, together with the consultants' conclusions, are summarised in [Figure 1](#).

QinetiQ was chosen through a publicly tendered competition

- 1.9 An invitation to tender was published in the Official Journal of the European Community on 6 November 1998. Thirty firms expressed an interest, of which four were short listed and invited to submit fully worked up tenders. Two of the short listed firms withdrew in the course of this process; one considered that the rate of investment return was not acceptable and the other could not meet the delivery deadline. Fully worked up tenders were received from the remaining two bidders on 17 May 1999. One of the remaining two bidders subsequently withdrew, again citing reservations over whether the income generated by the site would give them a sufficiently quick return on the development costs they expected to incur. This left only the bid from QinetiQ⁴.

- 1.10 The Public Record Office invited QinetiQ in June 1999 to give a final presentation and carry out a demonstration of elements of the project. The presentation covered customer service, and commercial, technical, implementation and administrative issues. QinetiQ and its subcontractors demonstrated the procedures for scanning the microfilm images of the census returns and data capture from returns. They also presented the proposed website design, illustrating the structure of the proposed search facilities. As a result of this presentation and subsequent assurances by QinetiQ on the various requirements of the project, QinetiQ passed the tender board and was appointed as preferred bidder on 20 August 1999. The contract was based on the standard Office of Government Commerce model for information technology projects. It was outcome based, and the commercial and development risks were transferred to the contractor.

- 1.11 The key points of the contractual arrangements between the Public Record Office and QinetiQ are set out in [Figure 2 overleaf](#).

- 1.12 To deliver its responsibilities, QinetiQ made use of the following subcontractors:

- Amey IT Services with its subcontractor BT Ignite for hardware and communications infrastructure;
- Sopheon UK Ltd for software development; and
- Enterprise and Supply Services, (a division of HM Prison Service) for data transcription.

- 1.13 The Public Record Office's original intention was that the arrangement should be on the basis of a Private Finance Initiative with project risk transferred to the private sector. The project agreement signed on 19 January 2000 was made between the Keeper of Public Records and the Secretary of State for Defence acting through QinetiQ. The project risks were therefore transferred from one to another part of the public sector. Additionally, the contract involves no management charges or fees payable to the contractor by the Public Record Office. Instead, the contractor recovers its costs through retention of subsequent revenues generated under the contract.

⁴ At the time the contract was originally let, the Public Record Office was dealing with the Defence Evaluation and Research Agency (DERA). A major part of the business of DERA was transferred to a new group of companies called QinetiQ plc in July 2001. The census contract was transferred to a company within the group called QinetiQ Ltd. 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 the 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.

1 Options rejected by the Public Sector Comparator review of the 1901 census

The Public Record Office's consultants commented on six other options of providing public access to the 1901 census.

Option	Conclusion
1. Manual release of census at Family Records Centre in Islington.	Did not contribute towards the long-term strategic goal of providing census data in electronic format.
2. Manual release of census at the Public Record Office at Kew	Did not contribute towards the long-term strategic goal of providing census data in electronic format, although viable as a short-term contingency.
3. Move Family Records Centre to new location, with manual release of census at new location	A major capital project, not achievable by 2 January 2002. Moving away from the Office of National Statistics, which is co-located at the Family Records Centre and which provides complementary data on births, deaths and marriages, would also have had a negative impact on the level of service provided.
4. Manual release of census through local record offices	Did not contribute towards the long-term strategic goal of providing census data in electronic format and would seriously fragment the quality of customer service.
5. Digitise and make census available over Internet through in-house development and operation by the Public Record Office	Key disadvantage - lack of an adequate range of skills within the Public Record Office. Headcount restrictions would have severely limited the ability to recruit suitable staff. The use of a volunteer labour force to undertake transcription would have entailed finding and training such a pool of labour. Maintaining the confidentiality of the data prior to January 2002 would have been logistically challenging. Other disadvantages would be questionable quality control and an uncertain implementation schedule.
6. Digitise and make census available over Internet with development by a contractor and subsequent in-house operation by the Public Record Office	Expensive for the Public Record Office - would involve using contractors and consultants at market rates.

Source: National Audit Office summary of the Public Record Office's consultants' report

The project was planned to be self-financing

1.14 The investment and development costs for the project were incurred by QinetiQ. The initial source of this funding was from working capital reserves held by QinetiQ rather than specific public funding from the Ministry of Defence. The ongoing operating costs are also borne by QinetiQ. The Public Record Office only incurs its own costs relating to contract management and administration.

1.15 The Public Record Office and QinetiQ will share net revenues equally, once QinetiQ's agreed development costs have been recovered. Initially, the establishment of a separate Future Census Fund⁵ was envisaged, once QinetiQ's costs had been recovered. However, a simpler

arrangement whereby QinetiQ and the Public Record Office share net revenues equally once QinetiQ's agreed costs have been recovered was subsequently agreed by both parties to be more appropriate. Net revenues comprise the revenue generated by the website less an agreed amount, originally anticipated in the contract to be around £82,000 per month, to cover QinetiQ's operating costs. Net revenues received by the Public Record Office are to be used to fund further digitisation of other record collections.

1.16 Revenues are raised by charges levied for access to documents from the website. Initial access to the website is free, as is searching the index. Viewing transcribed data costs 50 pence for an individual and then 50 pence for a list of all other people in that person's household. Viewing a digital image of the census page costs 75 pence. Users may either pay

⁵ The Future Census Fund was designed to earmark funding for work on digitising future censuses and was part of a complex formula for sharing revenues arising from the 1901 census project between the Public Record Office and QinetiQ.

2 Contractual arrangements agreed between the Public Record Office and QinetiQ

The contract between the Public Record Office and QinetiQ divided responsibilities for the project and the marketing strategy, and ownership and risk between the two parties.

Public Record Office	QinetiQ
<p>Responsibilities</p> <p>Makes available at agreed times the 1901 census microfilms.</p>	<p>Designs, develops, implements, commissions, operates and maintains an online access service, including collection of revenue, for the 1901 census.</p>
<p>Ownership and risk</p> <p>Retains ownership and risk in census data and intellectual property rights in all material developed there from, e.g. multimedia and audio-visual.</p> <p>Shall have a non-exclusive, royalty-free and irrevocable licence to use new inventions, design process or other technology created in the course of developing the database, and 50 per cent share of profits arising from their commercial exploitation.</p> <p>Retains ownership of any domain name or other identifier for specific use or access to the database. Licenses QinetiQ to use the domain name solely for the purpose and duration of project agreement.</p>	<p>Retains ownership and risk in tangible property and software used for the provision of services.</p> <p>Retains ownership of new inventions, design processes or other technology created in the course of developing the database, and 50 per cent share of profits arising from their commercial exploitation.</p> <p>May use any domain name in relation to the services solely in developing and hosting the database.</p>
<p>Marketing strategy</p> <p>Jointly responsible with QinetiQ for marketing the online service. The Public Record Office takes the lead in some areas, and in other areas, QinetiQ. Who is responsible for doing what is set out in a marketing strategy agreed in June 2001.</p> <p>Remains main focus of information for market segments A and B (family, local and social historians), particularly in the pre-launch period.</p>	<p>Jointly responsible with the Public Record Office for marketing the online service. QinetiQ takes the lead in some areas, and in other areas, the Public Record Office. Who is responsible for doing what is set out in a marketing strategy agreed in June 2001.</p> <p>Investigates potential marketing opportunities in respect of market segments C and D (the general public and schools), particularly in the post-launch period.</p>

Source: National Audit Office

through the purchase of prepaid vouchers, or through using their credit or debit cards online. The latter is subject to a minimum charge of £5. Fees for all the services provided by the Public Record Office are covered by a statutory instrument.

A cutting edge service was planned to be delivered

1.17 The Public Record Office planned that members of the public would access copies of original records by searching for particular individuals, by location, institution and vessel, and obtain selected images. It envisaged that the site would be further enhanced to enable searching by reference to specific addresses (a facility that is now operational). The factors contributing to the cutting edge nature of the project included:

- full transcription from the original records;
- social inclusiveness fostered by open access to the system with no need to register, free searching of primary indexes, and the low cost of further access to data;
- the flexibility of being able to buy subscription vouchers off-line at Public Record Office sites, libraries and local record centres;
- the high quality of the images of original documents;
- the range of search facilities; and
- subsequent amendment of data for reported errors.

1.18 This service was planned against the background of other sites which offered related services. In its market research, the Public Record Office took account of sites such as:

- FamilySearch, a website run for the Genealogical Society of Utah, launched in May 1999 with 640 million personal records;
- the 1.7 million records of Commonwealth forces personnel who died in the First and Second World Wars made available online by the Commonwealth War Graves Commission; and
- the Ellis Island website, launched in April 2001, which provided access to 22 million individuals listed on the Ellis Island passenger manifests.

In comparison, the 1901 census contains around 32.5 million records.

1.19 The Public Record Office set up a Census Advisory Panel with the objective of representing the main user communities for census returns. Membership of the Panel was by invitation from the Public Record Office and was drawn from a number of bodies including the Federation of Family History Societies, the Society of Genealogists, the Guild of One-Name Studies, the Association of Genealogists and Record Agents, the British Association for Local History and the University of Essex. The Panel held its first meeting in January 2000 and thereafter met every other month. Its terms of reference included advice to the Public Record Office and QinetiQ on matters of concern to users and on user priorities for development of the service. The Panel also provided an independent assessment on specific elements of the 1901 service as they were developed. Examples of the range of issues raised by the Panel are shown in **Figure 3**.

1.20 From 5 September 2002, the Census Advisory Panel became the Public Record Office Online Services Advisory Panel, and its terms of reference were widened to include any proposed digitisation projects and online services.

3 Issues raised by the Census Advisory Panel

- The accuracy of transcription
- IT familiarity for users without computers and/or Internet access
- The cost of Internet access
- Access via the index (i.e. which index elements would be available)
- The transcription of Welsh place names
- The availability and pricing of microfiche¹
- The 1891 census pilot, particularly the quality of transcription, and the removal of the 'Head of Household' feature
- The procedure for claiming a refund where an incorrect image had been downloaded due to a transcription error

NOTE

1. Many members of the Census Advisory Panel felt that the price of microfiche at £2.60 was too high. However, this was determined by statutory instrument and set on a cost recovery basis.

Source: National Audit Office review of the Census Advisory Panel's papers

Part 2

Problems arising during implementation

2.1 This part of the Report explains the nature of the problems that arose as the project proceeded. It covers the transcription of the original census forms into a digital format that could be processed electronically, the design of the website, the marketing of the project and the problems that occurred when the site was initially launched in January 2002.

Initially, insufficient progress was made converting census records into a digital format

2.2 The task of data preparation - digitisation - was a two-stage process.

- First, the Public Record Office's microfilm records of the original handwritten documents had to be scanned to produce a digital image that could be stored electronically and accessed via the Internet.
- Second, the data had to be transcribed, under secure conditions, from the microfilm of the original handwritten forms into standard digital records which could be combined to form a computer database which could then be accessed and searched via the Internet.

Once these two processes were complete, linkages were established within the database between the transcribed data and the digital image of the original document.

2.3 QinetiQ subcontracted the transcription of the census returns to Enterprise and Supply Services, a division of HM Prison Service, which entered into agreements with individual prisons for the transcription to be completed by prisoners⁶. During 2000, the Public Record Office and QinetiQ identified that Enterprise and Supply Services was not making sufficient progress in transcribing the census returns - the Prison Service had

other priorities for prisoners' time, such as providing them with information technology skills to increase their likelihood of employment and reduce their chances of reoffending. In June 2001, a review of the project was carried out by a member of staff from the Public Record Office together with a representative from the Office of Government Commerce, in accordance with the Office of Government Commerce's Gateway Review⁷ guidelines. The review concluded that the project was under control and was able to proceed with arrangements to 'go live' in January 2002, but confirmed the risk to the timetable for data transcription already identified by the Public Record Office and QinetiQ.

2.4 The Public Record Office and QinetiQ made considerable efforts to resolve the problem with Enterprise and Supply Services. Despite consultation at ministerial level, it was established that Enterprise and Supply Services could not complete the data preparation to a quality standard and to a timetable that would not prejudice the rest of the project.

Additional resources were required to complete the digitisation process

2.5 Therefore, with the agreement of QinetiQ, Enterprise and Supply Services, and the Public Record Office, Enterprise and Supply Services subcontracted and paid for the services of additional transcribers to complete the digitisation process outlined in paragraph 2.2. Under its contract with QinetiQ, HM Prison Service received income of £1.1 million from QinetiQ. The cost of the subcontracted work was £2.9 million which was funded from efficiency savings and from underspends in other areas. The overall net costs to HM Prison Service was therefore £1.8 million.

⁶ QinetiQ's decision to award the contract for transcription to HM Prison Service means that the risks inherent in the contract remained in the public sector.
⁷ A Gateway Review is a review of a procurement project carried out at a key decision point by a team of experienced people, independent of the project team.

- 2.6 The Public Record Office provided advice to Enterprise and Supply Services on testing the quality of transcription as part of the tendering process and on the development of Service Level Agreements with the appointed transcribers. The new subcontractors were Hays Document Management (based in India and Sri Lanka) and Atos Origin (based in the United Kingdom). Hays Document Management eventually carried out 78 per cent of the transcription work. In order for Hays Document Management to be used as a subcontractor, the Public Record Office had to secure the Office of National Statistics' agreement for the census forms to be transcribed offshore, under secure conditions.
- 2.7 In the course of the Public Sector Comparator study, the DMW Group considered an alternative approach to transcription using volunteers from the genealogical community. This option was rejected because of the implications for training and managing volunteers within the time constraints of the project and the logistics of maintaining the confidentiality of the census before 2 January 2002 using a large and dispersed workforce.

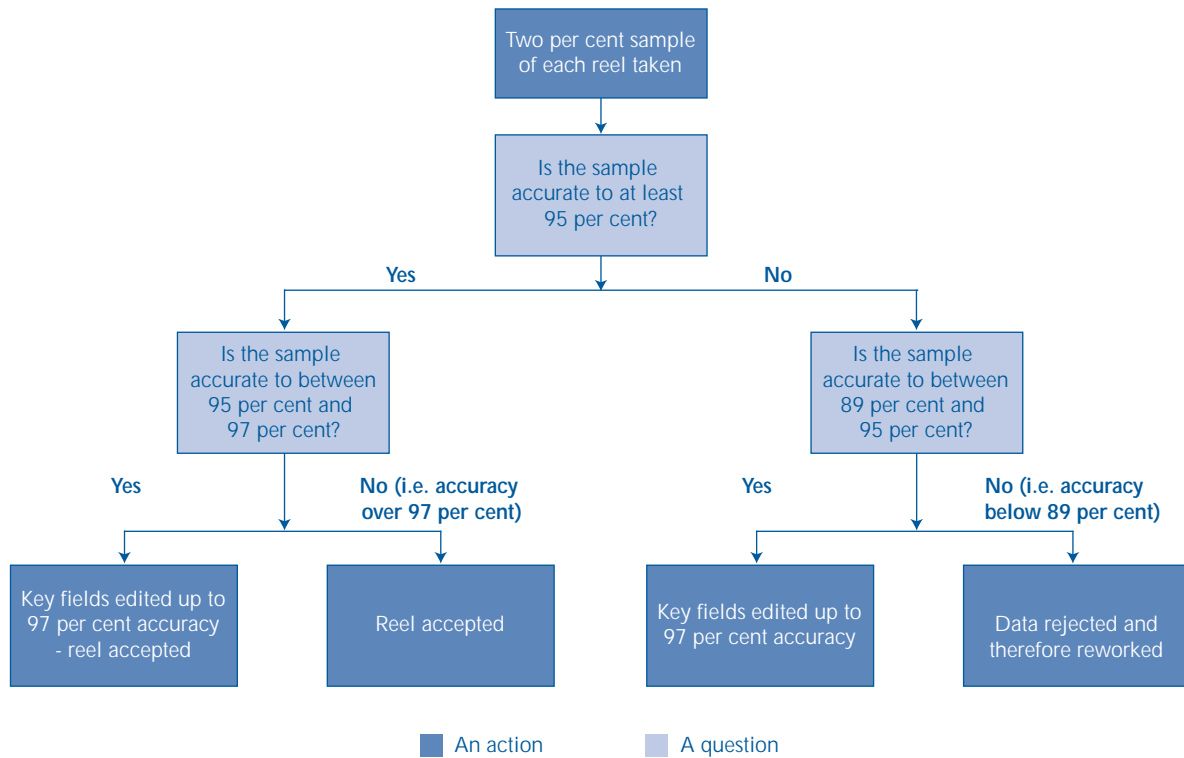
The Public Record Office stepped in to ensure quality control of data preparation

2.8 The major proportion of the work of quality checking the transcription of the 1901 census returns was undertaken by the Public Record Office. QinetiQ told the National Audit Office that it also committed additional resources to resolving the problems that arose on quality control of transcription. A sum of £350,000 was allocated by the Public Record Office to quality assure the transcribed data, with a further £500,000 of the quality assessment costs being met by QinetiQ. At the peak of the quality assurance work, the Public Record Office employed the equivalent of 68 full-time staff, over a period of approximately four months, on the process. Quality assurers were required to undertake and pass a test on reading late nineteenth century writing. In addition, ten staff from the Public Record Office with expertise in interpreting census enumeration returns spent between them 20 weeks in Sri Lanka and India to train local staff.

4 Transcription checks

Each reel of data was subject to two checks.

1. **A check for functionality and data integrity.**
If this was below the accuracy rate required, it was reworked.
2. **A check for accuracy of transcription.**



NOTE

After any necessary rework, of the 2,152 reels that were transcribed, 53 per cent were found to have an accuracy level of 95 per cent or above and were therefore accepted; the remaining 47 per cent were found to be between 89 and 94.9 per cent accurate.

Source: National Audit Office

This involved training sessions on the detailed transcription rules, the formation of the handwriting to be found on the returns and the etymology of Welsh place names. They also responded to queries raised by individual operators while transcription was taking place. Public Record Office staff ensured that keying operators had access to appropriate reference sources such as English and Welsh gazetteers and name listings. They also provided supervisors with detailed feedback on errors encountered during the quality assessment of the transcribed data so as to prevent such errors recurring.

2.9 The Public Record Office provided the reels of microfilm containing the data for digitisation, which QinetiQ converted onto compact disc. The transcription of each reel of data was subject to two checks (Figure 4).

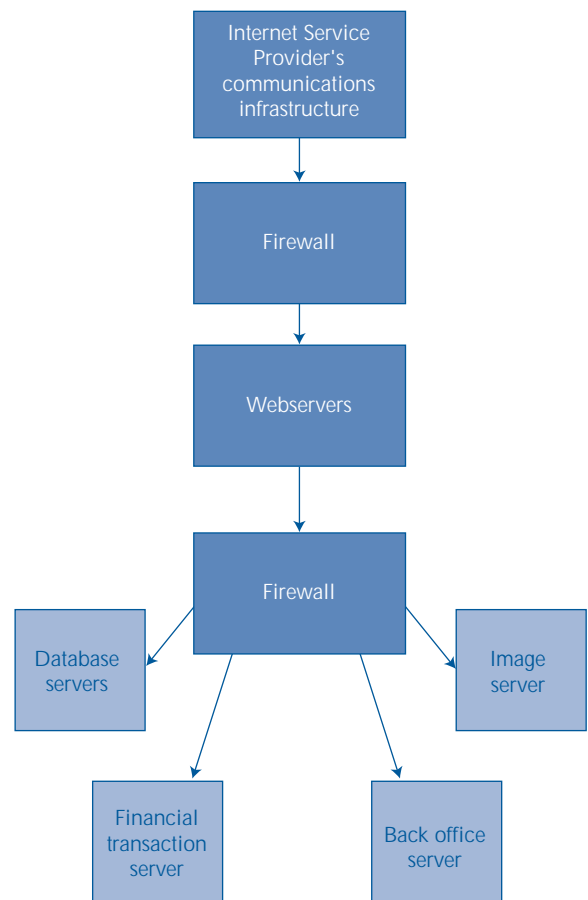
2.10 The Public Record Office and QinetiQ accepted the transcription task as complete by August 2001. However, they both also accepted that residual inaccuracies in transcription would remain due partly to the volume of data and also to the interpretation of Victorian handwriting. To enhance the quality of the data over time, the website provides a census amendment facility where users can report any errors that have been identified. However, errors or omissions that appear in the original census documents and which have been accurately transcribed are not classified as errors. Prior to the official relaunch on 21 November 2002, about 3,000 transcription errors had been reported by users, and 1,600 census data amendments had been dealt with by QinetiQ. New amendments are now dealt with by QinetiQ in blocks on at least a quarterly basis, as stated on the website.

The website design process was QinetiQ's responsibility

2.11 QinetiQ was responsible for the technical development of the website. The website is the package of hardware and software involved in providing an online service. This package includes the communications infrastructure which allows a user to log on to a webpage, the software links to what the user views on screen at the website, the database of information, the in-house information itself in the form of the database, the image archive, and other infrastructure to allow for processing financial transactions. The configuration of the site's hardware is set out in Figure 5.

2.12 The Public Record Office and QinetiQ decided to realign the management processes for the project in order to provide more effective and efficient management of the project, and to reflect the partnership

5 Configuration of the 1901 census website



Source: National Audit Office

nature of arrangements between them as regards the project as a whole. For example a Joint Programme Board was formed in February 2001, with membership drawn from both the Public Record Office and QinetiQ, under the chairmanship of the Keeper of Public Records.

2.13 To provide data on user profiles, QinetiQ and the Public Record Office ran a pilot project on 1891 census returns for the County of Norfolk. This pilot project was available to users from May to July 2001, and was scaled to provide access to 400,000 users per month. Usage statistics for the site showed that 187,000 users had accessed the site, considerably less than the available capacity. The pilot project was closed in August 2001 and QinetiQ used the results to enhance the website from the users' point of view. By September 2001, the Public Record Office and QinetiQ had agreed the modifications that should be made to the service.

A progress review concluded that the project should go ahead with its agreed launch date

2.14 The project control arrangements were assessed as part of the review of the project in June 2001. The review concluded that 'As a partnership arrangement, both sides are committed to the success of the project and operate under an "open book" arrangement'. The review stated that although some work was still needed on areas such as a more detailed implementation plan, the project should go ahead with its agreed launch date, and that the high level implementation plan, which set out key milestone dates, did not break down the activities into the level of detail that might be expected for the last six months of the programme. This detail was subsequently set out by QinetiQ for review at each meeting of the Project Board.

2.15 The National Audit Office published a report on risk management in August 2000 - Supporting innovation: Managing risk in government departments⁸. In this report, the Public Record Office's 1901 census project was used as a case study to illustrate good practice in risk recognition and management. At that time the contingency measures identified for the project were given as examples of effective measures to counter the identified risks. These measures principally related to safeguarding the Public Record Office's statutory commitment to make the census records public by having the returns available in microform, should the Internet access fail, and the use of the 1891 pilot study to assess potential user profile for the 1901 data.

The Joint Programme Board received regular reports on the project's progress

2.16 During the final stages of data loading, a technical problem was uncovered by Sopheon, QinetiQ's software subcontractor. The database was not operating as efficiently as required in relation to the performance of some of the indexes. The software contractor had assumed that the current version of the software in use could fully support all the manipulation of indexes required to provide the planned service. This problem related to the volume of activity being handled by the system and, whilst the software suppliers were aware of the issue, it was not formally referred to in the technical manual supporting the software. Having identified this problem, a temporary solution was put in place to remedy it. This problem further impeded pre-launch

testing. Data loading was completed by early December 2001, leaving only one month for QinetiQ to carry out pre-launch testing.

2.17 In the months leading up to the launch in January 2002, the Joint Programme Board received regular reports of progress on the data loading process, on testing of the system, and on checks on completeness. Any issues which remained were reviewed in detail by the Joint Programme Board. On the basis of the information received by the Joint Programme Board, which stated that the site was sufficiently robust to provide the planned service, the decision was made to launch the website on 2 January 2002.

The Public Record Office and QinetiQ agreed a joint marketing strategy for the census website

2.18 The Public Record Office and QinetiQ agreed a marketing strategy in June 2001 through which they were jointly responsible for marketing the census site. Four market segments were identified and targeted.

- United Kingdom family historians (this segment has been estimated at 250,000, based on membership of family history societies and groups).
- Local and social historians (including 200 universities, colleges and research councils in the higher education sector, and around 600 further education colleges).
- General public (including overseas users and professionals utilising search facilities).
- Educational users (including primary, secondary, sixth form and adult education).

2.19 The Public Record Office had significant information about family and social historians groups, traditionally the main users of census returns. It was, however, in the lead up to the launch, still collating information on potential demand from the wider public and users from the education sector. To reduce possible impact at the Family Records Centre, historians groups were to be targeted during the pre-launch period, in response to queries and issues raised in communications with the Public Record Office. Public talks were given at County Record Offices and libraries, but not at individual family history societies, who instead were encouraged to liaise with the appropriate County Record Office over hosting joint events. A questionnaire was sent out to a number of institutions who had expressed interest in providing access to the online service and in purchasing microfiche.

2.20 Some marketing was also carried out as part of the 1891 census pilot - for example, a live broadcast on BBC Radio Norfolk, and attendance at the Society of Genealogists' Family History Fairs in London, Yorkshire and Weston-super-Mare. The Public Record Office also promoted the 1901 census at the spring 2001 conference of the National Genealogical Society in Portland, Oregon, USA. A meeting was held for service centre representatives on 18 October 2001 to discuss the provision of the online service at local level. In December 2001, a 'How to Use' booklet was made available, and a CD training pack and instructions for purchasing vouchers were sent to service centres (those local record offices and libraries offering the online service). About 80 per cent of County Record Offices and almost 100 per cent of libraries had expressed an interest in making the online service available.

2.21 Market segments covering the wider public and the education sector were to be targeted after launch. It was anticipated that after the initial high level of demand, the number of hits on the site would stabilise. At this stage, the wider public and education users would be actively targeted in a bid to boost demand and halt any potential decline in interest. It was planned that QinetiQ would investigate potential marketing opportunities and undertake business case analysis to establish the financial viability of such plans.

2.22 The Public Record Office's intention was to have a low key launch, as reflected by its pre-launch strategy to target family and social historians groups, which were relatively small in comparison to the wider public and educational users. This approach was agreed by the Joint Programme Board.

2.23 Prior to the launch, the Public Record Office received a request from the Daily Telegraph for a press release. In response, a press release was issued and the Daily Telegraph ran an article on the 1901 census on 1 January 2002, the eve of the launch. Despite the Public Record Office's intention to have a low key launch, the article was the precursor to intensive and unexpected levels of press coverage in the days immediately after the launch: the website featured regularly on news bulletins and was covered by newspapers in, for example, England, Wales, and Scotland. During this period, staff at the Public Record Office received calls from around the United Kingdom and from other countries, such as South Africa, New Zealand and the United States of America.

When the 1901 census website was launched it became overwhelmed by a high level of demand from users

2.24 The 1901 census website was officially launched at 9 am on 2 January 2002. The service was scaled to provide access to 1 million users with a peak of 1.2 million users in a 24 hour period. However, by midday, 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. The excess demand of 24 times specified levels resulted in a number of technical consequences.

- The Internet Service Provider had to deploy a wider than expected bandwidth.
- Firewalls within the Internet service provider ceased functioning, affecting other customers' sites. Within the configuration, the 1901 census site did not have dedicated firewalls. Instead it shared its firewalls with other customers of the Internet Service Provider, particularly financial institutions with high volumes of transactions.
- The database engine was working close to capacity, even though prior tests had shown that it was more than capable of meeting design capacity of a peak of 1.2 million users over a 24 hour period.

Experience for most users was therefore poor whilst the site was overloaded. People were unable to access the system, typically receiving 'Page cannot be displayed' error messages on their computers.

2.25 On 7 January 2002, the Public Record Office and QinetiQ agreed to close the site to general Internet access to allow QinetiQ to undertake a technical investigation. Availability was then limited to online access at the Kew site, the Family Records Centre in Islington and a number of service centres throughout England and Wales. This availability persisted until 6 February 2002, when the Public Record Office agreed that QinetiQ could close the site to allow for further work and testing.

2.26 The Public Record Office and QinetiQ's calculations of planned capacity of 1.2 million users over a 24 hour period were based on independent market research into other genealogical sites offering similar services, and the results of their pilot project which contained 1891 census returns for the County of Norfolk. The 1891 pilot was undertaken in order to inform the 1901 census project, as the fields on the 1891 census returns were very similar to those on the 1901 census returns. Statistics from the 1891 pilot showed that:

- 90 per cent of users had carried out personal searches;
- the most popular chargeable element was the image;
- 25 per cent of users had requested a full transcription;
- the majority of users were based in the United Kingdom; and
- there had been more requests for an A3 postal print of the image than expected.

2.27 While market research provided some indications about comparable sites, comparison was very difficult due to variations in scale, size and the nature of the specific records made available. However, the FamilySearch website run for the Genealogical Society of Utah, which was launched in May 1999 with 640 million personal records, received four times the expected demand and crashed at launch. After relaunch, the site was receiving usage well within the site capacity.

2.28 Following the system crash, the Public Record Office considered that the key issue to address was the system's ability to handle and divert demand in excess of the site's planned capacity. The Public Record Office accepted QinetiQ's argument that the investment that would have been required by QinetiQ to secure proportionately higher capacity, to cope with the high level of demand experienced in early January 2002, would have been prohibitive and higher than Public Record Office's limits for the project. QinetiQ agreed to complete further work on the system so that it was better able to cope with the high levels of demand seen on 2 January 2002, and operate at specified levels even in the face of excessive demand. As part of the further work on the site completed by QinetiQ after 2 January 2002, a second database server was added and the existing server up-graded.

Part 3

How the implementation problems were resolved

3.1 This part of the Report describes how the problems arising during implementation were resolved, including what happened after the system crashed and how the Public Record Office managed the contract with QinetiQ. The timing of the progress to full availability of the website is explained and user reaction to the service described. Additionally, the lessons learned for the future are set out. Finally, a case study highlights the similarities and differences between the Public Record Office's experience and that of the General Register Office for Scotland in putting 1901 census data online.

After the crash, both the Public Record Office and QinetiQ instigated technical reviews

3.2 In order to identify the causes of the crash, QinetiQ had detailed discussions with its subcontractors and short-term solutions were considered. These solutions included producing a 'text only' version of the site, the route used by the BBC on 11 September 2001 when its site faced overwhelming demand. QinetiQ and its subcontractors carried out an evaluation of the problems affecting the site, and commenced a programme of further work, including full testing of all aspects of the site.

3.3 In the days immediately following the crash, the Public Record Office discussed the position with the Office of the e-Envoy. While the Office of the e-Envoy does not usually provide assistance to departments and agencies experiencing problems with the provision of online services, it agreed to do so for the 1901 census project because of its high profile and following a specific request from the Keeper of Public Records⁹. The Public Record Office drew up its own test programme, and engaged independent consultants to assist in carrying it out. Testing

covered both functionality and service robustness from the standpoint of a typical user experience and simulated large volumes of users, similar to the levels seen on 2 January 2002. The tests were designed to check that the volume of users specified in the contract could use the site and that excess users would be diverted from the main site to an off line message. The site was made available to users at the service centres (local record offices and libraries) during the testing period.

3.4 The Office of the e-Envoy appointed Loudcloud, the organisation which manages the UK Online website¹⁰, to undertake an architectural performance assessment of the website. Loudcloud produced two reports, the first in January 2002 and the second, more detailed one, in March 2002. The purpose of the first report was to describe the 'As-Is' systems utilised by QinetiQ and to make technical recommendations on the existing systems. The second report described the architectural enhancements being designed and implemented by QinetiQ to address the website's performance issues. Loudcloud made a number of key observations in its reports.

- **System testing.** The appropriate level of performance testing with a well understood end user profile and testing goal had not been performed.
- **Project management.** QinetiQ, the prime contractor, could have been more directly involved in the deliverables of the project, especially with so many third party contractors involved.
- **Specifications.** Some of the project requirements could have been specified more precisely. For example, what exactly was meant by '1.2 million users per day', what is a 'user' of the site, and how many 'hits' and 'page views' (minimum, maximum and average) does a user generate?

⁹ The Public Record Office made initial contact with the Office of the e-Envoy on 22 January 2002. It was involved in the project on an approximately monthly basis until the site was relaunched in November 2002.

¹⁰ UK online - www.ukonline.gov.uk - was launched in September 2000 as the main government initiative to enable everyone to make the most of the Internet. It is the government portal site which diverts users to many different sites.

- **Technical knowledge.** The 1901 census database is extremely large with over 175 million records. Dealing with databases of this capacity, which are searched extensively, requires specific knowledge in many areas - for example, very strong software engineering skills, and hardware and infrastructure design, and support skills. Having all these technical skills available is a challenging task but they are essential for a complex project of this magnitude.

The Public Record Office and QinetiQ reached agreement on system changes with the assistance of the Office of the e-Envoy

3.5 The Public Record Office and QinetiQ took immediate action, as soon as it became apparent that the website was experiencing problems. A meeting between the Keeper of Public Records and the Chief Executive of QinetiQ took place as early as 4 January 2002. The management of both the Public Record Office and QinetiQ were in constant daily contact in order to keep the Public Record Office informed of plans to restore the system to full operational use as soon as was practicable. The Public Record Office considered invoking potential sanctions in the contract. On balance the Public Record Office considered its interests would best be served by ensuring that QinetiQ delivered a working website tested to its own satisfaction. Accordingly, the Public Record Office gave assurances to the Lord Chancellor that the website would only be relaunched once the Public Record Office had independently tested the system handed back by QinetiQ.

3.6 At the end of April 2002, the website was handed over for the Public Record Office to undertake its testing programme carried out by external contractors. By this stage QinetiQ had carried out the following work.

- The database server capacity was increased and the transaction and back office servers were re-engineered to cope with this revised configuration.
- The bandwidth of the communications access was increased.
- Dedicated firewalls were installed.
- Load balancing was made more robust.
- Capacity for an off line message service - to which users in excess of capacity would be diverted - was increased to fifty million users per day.
- The software was revised and tested to accommodate the above changes to the site architecture.

- Software was developed to measure the level of access to the site by unregistered users (this was required due to the unusually open access to the site through which users were not required to register before having access to free search facilities).

3.7 QinetiQ told the National Audit Office that the further work outlined above was additional to the specifications of the contract between it and the Public Record Office, and was introduced in a spirit of partnership with the Public Record Office on the basis that the site should be proof against a repeat of the exceptional demand experienced at the original launch. On the basis of these considerations some of the changes were applied on a short term basis so that the capacity of the site can be reduced if it proves to be over specified for the level of ongoing demand after relaunch. The Public Record Office told the National Audit Office that it was initially not privy to much of the technical detail of the work undertaken by QinetiQ after 2 January 2002 until both the Public Record Office and QinetiQ had sight of the second independent report commissioned by the Office of the e-Envoy and produced by Loudcloud in March 2002 (see paragraph 3.4), which described the architectural changes being designed and implemented by QinetiQ to address the website's performance issues.

3.8 The Public Record Office's functionality testing revealed 22 issues of varying severity, which were classified as high (3), medium (7) or low (12) severity. In May 2002, testing by 100 users, mainly Public Record Office staff, over a period of two hours, showed that 74 per cent of all users had problems with the site, particularly with image viewing, error messages and printing of images. Keynote, the consultants appointed by the Public Record Office to assist with the test programme, carried out a load test, which simulated 100,000 users over two hours. Again, there were problems including very slow page response, limitation on connections, and a high rate of some error messages. The site was handed back to QinetiQ for these problems to be addressed.

3.9 QinetiQ acknowledges that a number of issues were addressed and resolved in May and June 2002, but a number of problems related to causes external to the site. Once these issues were addressed, QinetiQ maintain that no significant technical issues remained and that this was the position reached by 18 June 2002. The Public Record Office, based on its own testing carried out in late May and in June 2002, maintained that the site was still not performing effectively, and the Public Record Office asked QinetiQ to resolve the remaining technical problems.

3.10 As the time taken to complete the further work on the site lengthened, the Public Record Office, in June 2002, initiated contractual legal remedies. QinetiQ reserved its legal position. The Public Record Office, with the agreement of the Chief Executive of QinetiQ, requested that the Office of the e-Envoy carry out an independent assessment of the website testing strategy in order to resolve differences in the agreed way forward. The Office of the e-Envoy commissioned Loudcloud and another independent consultant to carry out the assessment. They issued their report, the third commissioned by the Office of the e-Envoy, in July 2002.

3.11 The independent consultants acting for the Office of the e-Envoy listed three options in their report.

- **Option 1:** Start public testing at once with the flow of access to the site set low and work on changes for release at the end of public testing with no further testing. This option carried the risk of database overload and adverse publicity.
- **Option 2:** Make some changes immediately, then test, followed by a period of public testing. This would cause further delay before the service was restored.
- **Option 3:** Launch without public testing (not recommended).

3.12 The Public Record Office was in favour of the second option, whereas QinetiQ was in favour of the first. Eventually, with the assistance of mediation by the Office of the e-Envoy, agreement was reached on a detailed action plan to implement the second option.

The website became fully operational on 21 November 2002

3.13 Following the agreement reached with the help of the Office of the e-Envoy the site was relaunched on a test basis on 6 August 2002, with availability restricted to the Family Records Centre and Kew during normal opening hours. On 27 August 2002, it was made available worldwide via the Internet, initially with restricted hours which were gradually extended (by 8 October, the site was publicly available between 9 am and 5 pm daily) until 21 November 2002 when Internet access was fully restored, 24 hours a day, 7 days a week. The facilities offered were as originally planned except for the address search enhancement which was agreed to be deferred but which has since been implemented.

The Public Record Office and QinetiQ shared the costs of putting things right

3.14 The costs recoverable by QinetiQ under the terms of the contract amount to £8.15 million, increased from an initial amount of £7.5 million. These costs include £0.4 million in respect of the 1891 census pilot. We were told by QinetiQ that it has spent further sums on the work completed during 2002 and that it has lodged a claim with the Public Record Office for a corresponding increase in the agreed development and operational costs. The Public Record Office has advised QinetiQ that it does not consider any of the additional expenditure to be for work that goes beyond that covered by the contractually agreed capped development costs. QinetiQ states that the additional costs it incurred during 2002 are related to enhancements of which the Public Record Office was fully aware. The Public Record Office incurred total net costs of £1.7 million, after taking account of £500,000 recovered from QinetiQ in respect of data transcription costs (which is the major part of the increase of £650,000 in QinetiQ's agreed costs recoverable under the contract).

3.15 By 31 October 2003 the cumulative usage of the site had grown to 61.3 million services delivered, comprising 54.7 million free-of-charge services and 6.6 million paid services. Gross revenues to 31 October 2003 total some £4.5 million. These figures were supplied by QinetiQ, which has retained all these revenues to cover its development and site operational costs. QinetiQ told us that site operational costs had yet to develop a monthly trend.

3.16 The revenues are collected and administered by QinetiQ. It will be important for the Public Record Office to obtain independent assurance on operating costs incurred by QinetiQ and the amounts of gross revenues received.

The relaunched website has been well received

3.17 The immediate response to the launch failure was one of great disappointment. Some members of the Census Advisory Panel felt that their warnings about potential high numbers had not been taken into account; others felt that no one could have foreseen what was going to happen on launch and they did not blame the Public Record Office or QinetiQ for lack of foresight. The Panel also enquired as to the causes of failure, the technical enhancements made to the site, the robustness of the search engine, and whether taxpayers' money was being used to finance the enhancements. Ongoing issues included the accuracy of transcription and the speed of error correction. The Public Record Office confirmed that compensation would only be paid where a transcription error had caused a user to download the wrong image.

3.18 Panel members became concerned at the time taken for QinetiQ to complete its further work on the site. Concerns tabled at the January 2003 meeting of the Online Services Advisory Panel centred on the implementation of an address search facility (now available to users), the best use of the wild card search facility, and the correction of identified transcription errors.

3.19 Since its relaunch, between 8,000 and 10,000 users visit the website daily. To establish what users think of the website since its restoration, we examined a small sample of users' comments forwarded to the Public Record Office, as well as comments directly recorded on genealogical websites. Comments were mainly from UK users, although there were comments from overseas users, including Poland, Switzerland and Australia.

3.20 The overall reaction from users has been one of delight mixed with relief. Users felt that it was well worth the wait, particularly those users who would find it difficult or costly to travel to London. The free element of the services was greatly appreciated, and there were few adverse comments about the payment of charges. Response times were judged not to be slow in comparison with other websites. One user, who reported a transcription error and received a reply within 36 hours, greatly appreciated the fact that it was not an automated response. Another user successfully traced a family on the 1891 census through to the 1901 census. The facilities on the website enabled the search to be completed in minutes, whereas the equivalent search on the 1891 census took several days.

3.21 Neither QinetiQ nor the Public Record Office have carried out a survey of users. No data is therefore currently available on the distribution of users of the online service by type of user (family historian, academic, member of the public, and so on) or by demographic indicator. It is not therefore possible to assess whether online access to the 1901 census has met its objective to widen social inclusion.

3.22 The National Audit Office contacted representatives from the Association of Genealogists and Researchers in Archives, the Society of Genealogists, and the academic community to ascertain their views on the operation of the website. **Figure 6** summarises some of their views.

6 Users' views on the website

Users of the website were broadly positive about the service provided by the website.

- 'Our overall view is that... the service is good but it could do better.'
- 'When the service works, it is excellent.'
- 'The online service is good in the sense of servicing the core constituency [but] poor in the sense of serving the constituency outside genealogy.'
- 'The needs of academics, local historians and demographers that are not looking for information by name but by place, occupations etc. have not yet been fully met.'
- 'The quality of transcription is only fair to good, owing to the number of inaccuracies in the data set.'
- On the quality of transcription 'this is clearly where the project falls down. I have been shown entries where the quality control has clearly been unsatisfactory. So I must judge transcription as poor.'
- 'The index online has made the census available to far more people than would previously been possible.'
- 'The website has met the needs of genealogists and researchers in archives. It has also benefited those who do not have physical access to the Family Records Centre (located in Islington), the National Archives (located at Kew), local record repositories, and so on.'
- 'Social inclusion has improved but only as far as genealogy and family history is concerned.'
- 'The quality of the scanned image is excellent and search facilities are generally good.'
- 'The search facilities are good - but they could be better. The index search results should show second forenames where they exist. It should be easier to search on address than it is.'
- 'The website help facilities and background information are good.'

Source: National Audit Office review of users' views on the website's operation

Making the 1901 census available online has implications for future projects

- 3.23 The level of change in use of the Internet and communications technology generally means that no clear prediction can be made of the specific circumstances that the Public Record Office will face when it comes to make the 1911 census available to the public. The Public Record Office has also developed new ways to exploit commercially public records. This section looks at the lessons to be learned for the future from the 1901 experience, taking account of the foregoing factors.
- 3.24 The project agreement provides that, subject to income becoming available and the continuation of a satisfactory service by QinetiQ, the 1841 to 1891 and the 1911 censuses should form a future part of the agreement between the Public Record Office and QinetiQ. Following the success of the 1891 census pilot, a paper for taking forward the digitisation of the 1881 and 1891 censuses was presented to the Joint Programme Board in September 2001. The paper was jointly agreed, at that stage, as a statement of intent. The project has not been taken forward subsequently. Changes in Crown Copyright and the implementation of the HM Treasury Policy Document entitled "Selling Services into the Wider Markets" mean that alternative licence arrangements, described further below, allow greater flexibility for commercial exploitation of public records than a more conventional PFI contract.
- 3.25 The Public Record Office licenses a wide range of applicants to use images of its holdings on a non-exclusive basis¹¹ as part of their commercial products or services, in return for royalty payments¹². Such licensing contracts do not constitute any form of partnership between that organisation and the Public Record Office. MyFamily.com has already applied for, and received, a non-exclusive licence to digitise the census returns from 1841 to 1901 and is in the process of making this material available on the Internet.
- 3.26 In November 2002, the Public Record Office announced a new opportunity, Licensed Internet Associateships, which will enable potential partners to participate in the commercial digitisation of genealogical and other records. The difference between a Licensed Internet Associateship and a non-exclusive licence is that an Associate will be actively associated with the Public Record Office. To be approved as a Licensed Internet Associate, an organisation needs to have a demonstrable and proven business track record, and must meet the minimum criteria of the Public Record Office's ethical standards. In return, the organisation receives support and expertise from the Public Record Office as well as being able to use the Public Record Office's branding and logo, including links to the Public Record Office website and online catalogue.
- 3.27 The Licensed Internet Associateship has been described as a halfway house between a public-private partnership and a non-exclusive licence. Being associated with the Public Record Office brand would bring commercial value, market expertise and document knowledge. This opportunity is considered more suitable for some of the larger collections (such as returns from censuses completed before 1901 and detailed records of soldiers who fought in the 1914-1918 war) which the Public Record Office could not digitise on its own through lack of funds. Using the revenue generated from Licensed Internet Associateships, the Public Record Office would digitise less-commercial family history collections or smaller material for inclusion in its own online image delivery system.
- 3.28 The Public Record Office considers that, as yet, it is too early to anticipate how the 1911 census will be made available to the public. The original census material has not yet been 'microfilmed' and it remains an open question as to whether this should be done at all, as technical developments may allow for the data to be scanned directly. Also, the Public Record Office may not necessarily carry out this process itself as has been the case in the past. More widely, future technical developments in access to and use of the Internet may have a significant bearing on what opportunities may prove feasible in dealing with the 1911 census.
- 3.29 Accepting the above uncertainties over the precise form of public access to the 1911 census, the following issues arose from the experience of making the 1901 census available online.
- The specific point at which census records become open to public access leads to an element of anticipation and pent up demand for the new information that they contain. This points towards a blanket rather than gradual release of the information. This approach, however, has unavoidable risks for the Public Record Office's reputation, which it is not possible to transfer to a third party.
 - The capacity required to cope with initial demand may far exceed the capacity required on an ongoing basis.

¹¹ A non-exclusive licence allows the holder to use images from the Public Record Office's holdings commercially in return for a royalty payment. The Public Record Office can issue any number of these licences - individual licensees do not, therefore, have sole use of the image(s) covered by the licence.

¹² While the detail of each contract may vary to ensure that the way the royalty payment is levied is tailored to the licensee's specific charging model, the rate at which the royalty is charged is seven per cent of net revenue.

- The requirement to go public on a specific day means that preparation of the data and development of the system has to be carried forward on a confidential basis. This requirement imposes considerable constraints compared to open access to data already in the public domain under the licensing arrangements described above.
- The ability to test directly user reaction to the actual data on a pilot basis is constrained due to the need to maintain confidentiality until the information is in the public domain. This constraint may have considerable consequences where the nature of the records changes significantly between one census and the next. However, the increasing availability of genealogical sites containing census information and the developing experience of the 1901 census should provide sounder data on which to estimate likely demand and user profile in relation to any future release online of new census data.
- Earlier completion of service development is needed to allow for independent assessment of the system pre-launch. This should reduce the risk of unanticipated consequences after the launch.
- Advance agreement on a post-launch disaster contingency plan would optimise the chances of rapid recovery from any future system failures arising at launch.
- The financial return from the 1901 census website is a factor that needs to be considered in deciding the precise contractual or licensing arrangement to be used for the launch of the 1911 census.

Scottish 1901 census data online

- 3.30 The arrangements through which the Scottish census data for 1901 were put on the Internet are summarised in Appendix 4. The following case study highlights the points of contrast between the experience of putting English and Scottish 1901 census data online.
- 3.31 This case study illustrates the very different circumstances surrounding the projects to place the English and Welsh 1901 census and the Scottish 1901 census online. 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.

CASE STUDY: PUTTING 1901 CENSUS DATA ONLINE IN ENGLAND AND WALES AND IN SCOTLAND

Similarities

- External contractors used to develop and operate websites at their own cost.
- Revenues generated from the website shared with contractors.
- Digitisation process was undertaken overseas.
- Similar objectives - improved quality of service and contributing to the modernising government agenda.

Differences

- Statutory requirement to make the England and Wales census available to the public on 2 January 2002. There is no statutory requirement to make the Scottish census data available to the public on or by a specific date.
- Scotland has only around 244,000 images at the level of enumerator's summary for a population of 4.5 million compared to 1.42 million images of individual returns for England and Wales for a population of 32.5 million.
- The Scottish image data is indexed but the information contained on the images of the enumerators' summaries was not separately transcribed.
- The Scottish site received 123,000 hits during the first weekend; England and Wales received 29 million hits during the first week. Demand settled at 8,000 - 9,000 accesses per month for the Scottish site in its then current form up to the end of August 2002. Demand for the English site is 8,000-10,000 users per day (equating to approximately 243,000 - 304,000 users per month).
- 1891 census data for Scotland, along with indexes to other material, was already available with a site charge paid by users to gain access, thus contributing to a climate of chargeable access to on-site and online data for Scotland at the time the 1901 census went online. Potential users of the English site had no such precedent.
- The Scottish site charges a fee in order to get access. The charge at launch was £6 for a 24 hour period, it is now £6 for a 48 hour period. In England, initial access and certain searches are free. Charges are levied for specific searches with charges between 50p and 75p for viewing documents, with a minimum charge of £5 when a credit card is used to pay online.
- Contractors used in Scotland were wholly within the private sector and had experience in working with Scottish records since 1997. In England the contractor was originally a public sector entity which had no previous experience of working with the Public Record Office, but did have extensive expertise in Internet systems.

Appendix 1

Our approach to the study

- 1 This report examines the basis for the project and the reasons for the delay in restoring services to the public. It looks at lessons learned of general relevance to delivery of central government services across the Internet.
- 2 We considered:
 - the reasons why the project was contracted out and the contracting process;
 - the planned nature of the service to be provided by the website;
 - the way the project was financed;
 - the record transcription and website design process;
 - the pre-launch marketing considerations;
 - the extent of testing programmes;
 - the reasons for the initial crash of the system;
 - the approach to restoring the system;
 - the relationship between the Public Record Office and QinetiQ and the role of the Office of the e-Envoy;
 - user reaction to the website, usage levels of the website and revenues generated; and
 - implications for future release of other census data and wider lessons for managing online projects in the public sector.

- 3 The key methods we used were as follows.

- We undertook a detailed examination of contract documentation, project management records, independent reports commissioned by the Public Record Office, and correspondence between key parties - QinetiQ, the Public Record Office, and independent advisors. The key documents we examined are listed in the table below.

Key documents examined during the course of our examination

- The contract between the Public Record Office and QinetiQ
 - The Public Sector Comparator report
 - A review of the project in accordance with the Office of Government Commerce's Gateway Review guidelines
 - Transcription Project Board minutes
 - Marketing plans
 - Loudcloud reports on the project status
 - Project Board Minutes
 - Census Advisory Panel and Public Record Office Online Services Advisory Panel minutes
- We interviewed senior staff at the Public Record Office, the General Register Office for Scotland and QinetiQ.
 - We made direct contact with user groups to ascertain their views on the operation of the website following its relaunch.

Appendix 2

Chronology

Appointing QinetiQ	
June 1998	The Public Record Office considers the implications of making the 1901 census available digitally, and concludes that the project could be managed using Private Finance Initiative arrangements.
6 November 1998	The Public Record Office publishes an invitation to tender for the project in the Official Journal of the European Community. Thirty companies showed an interest - four were shortlisted and invited to submit fully worked up tenders.
May 1999	The Public Record Office commissions consultants to complete a public sector comparator exercise. The consultants conclude that the Private Finance Initiative should be adopted.
20 August 1999	The Public Record Office appoints QinetiQ as the preferred bidder for the project.
19 January 2000	The Public Record Office and QinetiQ sign the project agreement.
January 2000	The Census Advisory Panel holds its first meeting (and thereafter meets bi-monthly).
Developing the online service	
May-August 2001	The Public Record Office and QinetiQ run a pilot project based around the 1891 census returns for the County of Norfolk to provide data on user profiles.
June 2001	A team comprising a member of staff from the Public Record Office and a representative from the Office of Government Commerce completes a review of the project. The review concludes that the project is under control and that it could proceed with arrangements to 'go live' in January 2002, but confirms the risk to the timetable for data transcription already identified by the Public Record Office and QinetiQ. The Public Record Office and QinetiQ agree a marketing strategy.
November 2001	QinetiQ and its subcontractors complete data loading.
The launch of the online service	
2 January 2002	The 1901 census website is officially launched by the Public Record Office at 9 am By midday, the site becomes overwhelmed with users. Access to the website is restricted to the Public Record Office at Kew, the Family Records Centre in Islington, and some service centres.
3-6 January 2002	The site continues to receive a very high level of users, overwhelming the site.
4 January 2002	The Keeper of Public Records and the Chief Executive of QinetiQ meet to discuss the problems with the website.

The launch of the online service (continued)	
7 January-6 February 2002	The Public Record Office and QinetiQ close the site to general Internet access to allow QinetiQ to undertake a general technical investigation. Availability is limited to online access at the Kew site, the Family Records Centre in Islington, and a number of service centres.
27 January 2002	Loudcloud, a consultant appointed by the Office of the e-Envoy, publishes its first report - a technical analysis of the existing system, including recommendations for improvements.
6 February 2002	The Public Record Office agrees that QinetiQ can close the site for further work and testing.
8 February 2002	The Public Record Office commissions an independent test of the system prior to relaunch.
8 March 2002	Action plan and testing strategies developed for independent testing.
March 2002	Loudcloud publishes its second report, on the architectural enhancements being designed and implemented by QinetiQ to address the website's performance issues.
29 April-7 May 2002	QinetiQ hands over the site to the Public Record Office for testing. The Public Record Office's functionality testing reveals 22 issues of varying severity. One hundred users test the site over a period of two hours - 74 per cent encountered problems - for instance, with image viewing. The Public Record Office's consultants, Keynote, complete a load test which simulates 100,000 users over two hours. Problems encountered include slow response times and a high rate of error messages.
April 2002	The website is handed over to the Public Record Office so that its contractors can complete a testing programme.
18 May-3 July 2002	Further testing of improvements to the site.
June 2002	The Public Record Office, with the agreement of QinetiQ's Chief Executive, requested that the Office of the e-Envoy carry out an independent assessment of the website testing strategy.
4-31 July 2002	Further enhancements and testing agreed following a report from the Office of the e-Envoy.
July 2002	The Public Record Office and QinetiQ agree on one of the three options suggested in the report of the Office of the e-Envoy's independent consultants on a website testing strategy.
6-26 August 2002	The Public Record Office and the Family Records Office begin user testing, with availability restricted to the Public Record Office and the Family Records Centre during normal working hours. Various problems are encountered - such as the slow page response time and the unavailability of the site - and overcome.

The launch of the online service (continued)	
27 August 2002	The site is made available to the public on a restricted basis - Monday to Saturday between 9 am and 7 pm; access in the evening is limited.
28 August-2 October 2002	Various problems are encountered and the site is shut down intermittently while the problems are resolved.
5 September 2002	The Census Advisory Panel becomes the Public Record Office Online Services Advisory Panel.
8 October 2002	Access to the site is extended to between 9 am and 5 am daily.
14 October-15 November 2002	The site becomes unavailable on various dates for an average of two hours while problems are resolved. The site is up and running again on the same day.
21 November 2002	The site becomes fully operational with 24 hour access, seven days a week.
November 2002	The Public Record Office announces the introduction of Licensed Internet Associateships, which will enable potential partners to participate in the commercial digitisation of genealogical and other records.

Appendix 3

A glossary of terms

Bandwidth	Refers to information-carrying capacity and is usually measured in bytes or bits per second.
Cookie	A cookie is a mechanism that allows a website to store its own information about a user on the user's own computer. Typically, a cookie records the user's preferences when using a particular site.
Cookie limit	<p>Also called a "cookie counter". This acts as a gatekeeper, and its purpose is to ensure that a website is not flooded and brought down by too many requests. By assigning and counting the number of cookies issued over a moving window in time, the intention is to give preferential access to those users whilst they are in the middle of a transaction, or over a defined period of time.</p> <p>Once the preset limit has been reached, a mechanism will be activated to divert additional users to a different site. The gatekeeper will then release the mechanism once a defined number of cookies become available for reissue. The site to which those users who cannot be issued with a cookie are diverted, will have a message such as "The site is currently too busy, but please take a tour of our site instead".</p>
Digitisation	<p>This involves three elements:</p> <p>(a) scanning the microfilms of the original census returns and creating an electronic image of each page of the returns.</p> <p>(b) transcribing the information from the returns and creating a database with a searchable index.</p> <p>(c) making the images and database available over the Internet.</p>
Enumeration	The act of taking a census.
Enumerator	A person appointed to deliver and collect household schedules, assist with their completion and then copy out the information into the Enumeration Book.
Firewall	A combination of hardware and software designed to protect a networked server from damage (intentional or otherwise) by those who log on to it.
Gateway review	<p>A review of a procurement project carried out at a key decision point by a team of experienced people, independent of the project team. The Gateway process, as defined by the Office of Government Commerce, considers the project at critical points in its development, which are identified as Gateways. There are six Gateways during the life cycle of a project:</p> <p>Gateway 0: Strategic assessment</p> <p>Gateway 1: Business justification</p> <p>Gateway 2: Procurement strategy</p> <p>Gateway 3: Investment decision</p> <p>Gateway 4: Readiness for service</p> <p>Gateway 5: Benefits evaluation.</p>

Institution	Any public or charitable buildings such as prisons, hospitals, asylums, schools, colleges, and barracks. Institutions housing more than 100 people were usually detailed in distinct Institutional Books.
Internet Service Provider (ISP)	An organisation that provides access to the Internet and other related services such as website building.
Load balancer	Hardware or software devices that can detect when a server becomes unstable or unavailable, and control the flow of incoming traffic accordingly.
Outsourcing	The process by which functions performed by the organisation are contracted out for operation, on the organisation's behalf, by third parties.
Private Finance Initiative (PFI)	A government initiative where the public sector contracts with the private sector to purchase services on a long-term basis so as to take advantage of private sector management skills incentivised by having private finance at risk.
Public-private partnership (PPP)	An arrangement whereby, typically, the public sector and private sector form joint ventures to improve the efficiency of public sector operations. They can be classified in national accounts to either the public or private sectors depending on who has the most control over the PPP.
Server	Usually a computer that provides the information, files, web pages, and other services to the client that logs on to it. Also used to describe the software and operating system designed to run server hardware.
Service delivered	The principal services are searches, downloads of images and transcripts, and print requests.
Service Level Agreement	An agreement between parties, setting out in detail the level of service to be performed under a contract.
Vessel	Ships and boats including the Royal Navy and those engaged in trade, fishing, transporting passengers and inland navigation. The census returns include foreign vessels in British ports and Royal Navy vessels at sea and in foreign ports.

Appendix 4

Putting 1901 census data for Scotland online

1. Statutory basis

Under the provisions of the Census Act 1920, the General Register Office for Scotland is not required to release Scottish census data to the general public, allowing greater flexibility in making the data available. The Scottish census data was made available on microfilm from November 2001.

2. Objectives

The General Register Office for Scotland's wider strategy is to make the entire database of its historical records available to customers as a digital image, as well as on microfilm and microfiche. Launching the 1901 images online was part of this wider initiative, being a broad digitisation programme, namely the Digital Imaging of the Genealogical Records of Scotland's people.

As well as putting the records online, an objective was to provide an enhanced service in New Register House, making the service more efficient and easier to use. A further objective was complying with the e-government and modernising government agendas.

3. Basis of availability

Optical Media Systems Ltd entered into a five year contract with the General Register Office for Scotland in September 1997 to host a website for Scottish records. As part of this contract, the 1891 Scottish census data was put online in August 2001 and the 1901 Scottish census data was put online on 24 January 2002. Optical Media System's contract came to a natural end on 31 August 2002 and, following an open tendering process, was not renewed.

From 1 September 2002, Scotland On Line hold a five-year contract with the General Register Office for Scotland to provide customers with pay-per-view web access not only to searchable-text index and digital images of the 1891 and 1901 census pages as at present, but also to digital images of most of the statutory registers.

4. Scale

In the month of January 2002, during which the 1901 census images were released, 22,000 registered users visited the site and paid for access to it. For the remainder of the Optical Media Services Limited's contract period the number of users visiting the site and paying for access settled down to between 8,000 and 9,000 per month.

5. Cost

Putting the 1901 census online formed part of an overall digitisation programme for the General Register Office for Scotland. The overall cost of this programme is £3 million over three years, funded by the Scottish Executive. Specific costs for the 1901 census which can be readily identified include £120,000 to create the index, and £12,000 for digitising the images. There are some quality assurance costs incurred as part of the wider project. The cost for putting the data online has been borne by both contractors.

6. Size

The Scottish 1901 census has a population of 4.5 million. Digitisation was carried out at the level of enumerators' summaries, rather than individual returns, resulting in only 244,452 digital images.

7. Charging

The charge at the time of launch was £6 for up to 24 hours usage, which gave the user 30 credits. The charge is now £6 for a 48 hour period. It costs one credit to view the results of a search and six to download an image. It does not allow any free searches except for surnames, which just lists the number of people with that surname contained in the census and other records included on the site.

Revenues generated are allocated between the General Register Office for Scotland and the contractor based on commercial agreement between the two parties.

8. Demand

The Scots Origins site registered 123,000 hits on its homepage during the first weekend of the 1901 census images being available.

9. Timetable

The General Register Office for Scotland put in place a phased release of the records to control demand for the information contained within the census:

- November 2001: The General Register Office for Scotland began deliveries of the unindexed 35mm roll-microfilm to libraries, archives, and family-history societies who had purchased it.
- 3 December 2001: Customers in New Register House were able to view roll-microfilm of the 1901 census with a searchable electronic index.
- 24 January 2002: The census was launched online.

10. Digitisation process

The process for Scotland included:

- microfilm the enumeration books from the Scottish census of 31 March 1901 on 35mm archive film;
- scan the films to create bitonal digital images;
- index of the 1901 census, done offshore from the scanned images; and
- index entries linked to the digital images.

Extensive quality assurance checks were made on the images, index entries and linkages, both by the contractor and by a dedicated team within the General Register Office for Scotland.