Technical Papers

COMPUTERISATION OF THE INDEXES TO THE STATUTORY REGISTERS OF BIRTHS, DEATHS AND MARRIAGES IN SCOTLAND

TABLE OF CONTENTS

* * * * * * *

FOREWORD

This report presents a description of a project to computerize the indexes of births, deaths, and marriages in Scotland. As in many other countries, for some time computers have been used in the production and analysis of vital statistics. Like a number of other countries also, Scotland has more recently given major attention to the application of computer technology to the operational activities of civil registration. The author of this report provides an account of the planning and implementation of the Computerisation of Indexes project in the hope that it will give civil registration officials in other countries an idea of the nature and scale of the undertaking, and of some of the problems which are likely to arise.

The views expressed in this report are those of the author and do not necessarily reflect those of the IIVRS.

There are no restrictions on the use of materials published by IIVRS. Materials from this publication may be quoted or duplicated without permission.
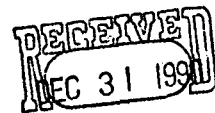
COMPUTERISATION OF THE INDEXES TO THE STATUTORY REGISTERS
OF BIRTHS, DEATHS, AND MARRIAGES IN SCOTLAND

by David Brownlee
Senior Executive Officer
General Register Office for Scotland

1.   BACKGROUND

1.1   General

1.1.1   The General Register Office for Scotland (GRO(S)) was established following an Act of the United Kingdom Parliament in 1854, to provide for the better registration of births, deaths and marraiges in Scotland, transferring the responsibility from church to state and putting a statutory obligation on individuals to register vital events.   Two of the current specific aims of the Department are:

    i.   to ensure that the records in the Registrar General's custody are kept in a satisfactory state of preservation, in security, and in a way which will enable the information to be recovered in the event of disaster;

    ii.   to provide a satisfactory service which enables the public to obtain access to and extracts or copies from the records in authorised circumstances.

1.1.2.   Since the mid-1960s GRO(S) have made increasing use of computers in the production and analysis of statistics, and currently are undertaking a project of converting the indexes to the statutory registers of births, deaths and marriages to computer-readable format so that they can be available to staff and customers 'on-line'.

1.1.3.   This paper gives an outline description of the indexes system as it existed prior to the introduction of computers, and gives an account of the planning and implementation of the Computerisation of Indexes project.   We hope it will give other Registration Authorities who may be thinking of similar projects in their countries an idea of the nature and scale of the undertaking, and of some of the problems which are likely to arise in the course of project development.

1.2   Description of the Indexes System prior to the Introduction of Computers

1.2.1   As mentioned above, one of the principal functions of the GRO(S) is to hold and make available for searching the statutory registers of births, deaths and marriages, which date from 1855. Under the 1965 Registration of Births, Deaths and Marriages (Scotland) Act and its predecessor the 1854 Act, the Registrar General is required to make and maintain alphabetical indexes to the entries in the registers in the GRO(S).   On payment of the prescribed fee, he must cause a search to be made in the indexes or allow a person to search in them and issue to any person an extract of any entry which that person may require.

1.2.2   The annual index is the main finding aid for locating individual entries in these registers. Paper-based indexes were subdivided into volumes by event type, year and sex, and within these divisions gave an alphabetic list of individuals, together with a Registration District (RD) and entry number for each event.   The RD and entry number were the keys to particular entries in the registers. Certain other information, such as mother's maiden surname, spouse's surname and age at death, were included to help distinguish between entries with common names.

1.2.3   The arrangement of the indexes determined the way in which they were used.   A typical user would know the event type, the sex and approximate year of the event.   This would narrow down the search to one volume for each possible year.   The user also had to know the surname (though not perhaps the exact spelling) and know either one or more of the forenames.   Unless the name was a common one, this was generally sufficient information to identify the particular entry.   In cases of common names, the other information mentioned in 1.2.2 could be used; in addition, the name of the registration district would help a user who knew in approximately which part of the country an event took place.

1.2.4　　GRO(S) staff searching in the indexes would typically have the information mentioned above, including the supposed year of the event.　The normal pattern of search would be to check the year given and 2 years on either side.　If the entry were not found within this period, a "no trace" return would be made.

1.2.5　　The indexes were in constant use by GRO(S) staff, with an estimated 114,000 accesses being made each year.　The indexes were also being heavily used by the public, making an estimated 400,000 accesses per year.　This gave a total annual figure of over half a million accesses to the indexes by public and staff.　In the mid-1980s there was every indication that genealogy was continuing to grow in popularity.

1.3　　Problems associated with Paper-based Indexes

1.3.1　　There were two major problems arising from the increasing use made of these indexes. The first was how to conserve the physical condition of the volumes and the second was how to cater for a steadily growing demand for access to them.

1.3.2　　Constant use was taking a heavy toll of the bindings and ultimately of the pages of these books.　Although some bindings had lasted 50 years or more, they were now being subjected to unprecedented levels of use.

1.3.3　　There was uncertainty about the remaining life of the index volumes.　Rebinding would certainly prolong it, but not indefinitely.　Eventually, the pages would become too badly damaged to be used further.　Restorative action had to be taken before this stage was reached.

1.3.4　　Increasing demand for searching threatened to swamp the existing index sets.　The indexes were one of the principal bottlenecks leading to searchers being turned away.　If the expected demand in the future were to be catered for, then one or more extra sets of indexes would have had to be provided.

1.4　　General User Requirement

1.4.1　　The Department required a system which would allow growing numbers of people to use the index but without slowing down access times and without causing disruption to register searching.

1.4.2　　Any system would have to allow for the expected rise in demand for searching.

1.4.3　　There had to be no significant increase in the time taken for searching, either because of the mechanics of searching or because there were insufficient facilities available.

1.4.4　　Indexes were being constantly updated because of re-registrations, change of name and correction of error.　These changes had to continue to be made within a maximum of one week from the time of notification by the section making the change.　All working copies of the index had to be amended at the same time.

1.4.5　　Steps to replace the paper indexes had not to interfere unduly with day to day work in the searching sections or with public access to the Index.

2.　　OPTIONS FOR REPLACEMENT OF THE PAPER-BASED SYSTEM

2.1　　Three options for replacing the paper-based indexes were considered.　The first option was microform and would have involved converting the paper index images into microfiche.　The second option was in fact to continue with paper, with a photocopy programme being set-up to make additional copies of the index available throughout the Office.　The third option was to convert all indexes to computer-readable format and make them available to staff and public via screen and keyboard devices.

2.2　　Microform

2.2.1　　The use of microfiche for index consultation had been extensively investigated over a number of years.　It had some potential advantages in view of our existing heavy commitment to filming the statutory registers and the consequent expertise and experience of our staff.　Microfiche would have been considerably cheaper than photocopying and this, together with its compact nature, would have enabled us to make extra sets.

2.2.2 Experiments with this medium, however, indicated serious drawbacks:

    i.   Size of paper sheets were such that it would have been impossible to film and then display a whole page in a standard viewer. Pages would either have had to be filmed twice or cut in half.

    ii.   It would have taken longer to identify and extract the correct fiche than to find a paper-based volume.

    iii.   There would be considerable risk of incorrect placement of fiche, especially if multiple sets were used in busy work areas.

    iv.   Because index searching would be slower with a fiche system, it would have been necessary to provide substantial numbers of microfiche viewers.

2.2.3 The conclusion of the examination of microfiche was that it would not be a suitable medium for replacing the indexes.

2.3 Photocopying

2.3.1 In this option, the existing index sheets, including manuscript amendments, would have been photocopied. A number of difficulties were envisaged with this option:

    i.   After copying, it would have been necessary to keep the original sheets and to continue to amend them, so that future copies could be made from them.

    ii.   The 1914-1965 indexes were on very large paper and the image was greater than standard A3-size paper. The cheapest way of dealing with them, therefore, would have been to guillotine the sheets, re-number and copy them at A3. This would have meant doubling the number of volumes.

    iii.   To ensure that the amendments, some of which were in blue ink, were properly reproduced, would have required careful checking of the copies before collation and binding.

    iv.   Copying would be one-sided which would have doubled the entire binding cost.

2.3.2 As well as the conservation of the indexes, the increasing public demand for access would have necessitated production of an extra set. As use of the indexes increased, more accommodation would have been needed and space was already at a premium.

2.3.3 A further disadvantage of the paper system was that it was not compatible with Department's medium and long-term strategy for records, that of gradually moving towards a system of electronic storage and access of registers.

2.4 A computerised Index

2.4.1 This option would establish a computer system in New Register House, holding index data on computer and driving a number of terminals throughout the building, for use by both GRO(S) staff and public. It was envisaged that there would be a phased take-on of data, with the numbers of terminals throughout the building increasing as more data became available on-line.

2.4.2 The computerised solution was identified as the most suitable replacement for the paper index. A sound Business Case, therefore, had to be made to the Government department responsible (HM Treasury) to obtain finance for the project. This was not easy given the considerable one-off costs of data conversion for such a project (an estimated 35 million records from 1855 to date) which would be incurred in the early years. We were able to convince our paymasters of the strategic and long-term benefits which would accrue well into the future from the project, however, and in late 1986 the project was given the go-ahead.

3. MANAGING THE PROJECT

3.1 General

In view of the nature and timescale of the overall project, it was decided at an early stage that planning and control would be more effectively achieved by splitting the project into a series of smaller sub-projects. These were as follows:

    i.   Procurement of computer hardware and software

    ii.   Data conversion

    iii.   Design and development of an on-line index retrieval system

    iv.   Accommodation matters

3.2   Data Conversion

3.2.1   Prior to the introduction of a computerised index, the indexes to the statutory registers of Scotland from 1855 to date were held in a variety of formats. Between 1855 and 1865 they were completed in manuscript; from 1865 to 1913 in letterpress format; from 1914 to 1965 in typescript format, and from 1965 to date the indexes were produced on computer as a by-product of the registration of vital events system, with a printed list being runoff and bound for subsequent reference by staff and public.

3.2.2   As previously mentioned, it was estimated that for the period 1855 to 1965 there were approximately 35 million index records which would have to be converted to computer-readable format. On average, it was calculated that each record contained 40 characters. Clearly, this represented a sizeable data conversion exercise using traditional keying methods. For this reason, an optical scanner was procured which could be 'trained' to recognise the type of characters being scanned. This appeared to be the most cost-effective method of data capture. In practice, however, the results achieved using intelligent character recognition did not live up to early promise. This was for the most part due to the poor quality of material being scanned. In older index volumes, problems with ink 'bleeding' into the paper, broken or badly-formed characters and letters joined together meant that the level of operator intervention was such that throughput speed was adversely affected.

3.2.3   When the original Business Case for the project was drawn up, a five-year timescale had been planned for the conversion of all historical indexes to computer-readable format, as other developments depended on this project finishing on time. The Project Board did not wish to see this timescale lengthened, so a further case had to be made for additional resources to allow the data conversion sub-project to stay on schedule for completion within five years. We were able to convince the relevant decision makers that the project was still a worthwhile one despite the additional one-off data capture costs; as the benefits from this project will accrue long into the future, a longer timescale was accepted in the investment appraisal to justify additional costs involved in manual keying during the early years of the project.

3.2.4   When paper index records are being keyed, the data fields name, first initial and entry number are also verified to cut down on possible error. Once data have been captured, they are checked for accuracy using a series of programs developed by out computer staff. It is impossible to ensure complete accuracy in an application such as this; for instance, if a surname is slightly misspelled but still appears in its correct alphabetical sequence, the error will pass undetected. This is serious for searchers because the record could 'disappear'. One very effective check which we carry out is referred to as the 'missing entry' check. In this program, each index record for each event type within a given year is sorted into entry number within registration district number; 'missing' numbers are listed and staff then check the original registers to find out what the correct entry should be. This check has revealed some errors in the original paper indexes, particularly where a birth has been re-registered some years after the original event. It is perhaps not surprising that errors and omissions in the original indexes should have passed undetected for so many years, as a sizeable percentage of index entries will possibly never be accessed. The problem, of course, is that one has no way of knowing which fall into this category!

3.2.5   Before embarking on the data capture sub-project, investigations were carried out into which type of registers and which period were most commonly accessed. The results of these investigations have determined the order in which indexes have been converted and taken onto the computer database. Data capture is scheduled for completion by March 1993 and the project is currently running within this timescale.

3.3   Procurement of Computer Hardware and Software

3.3.1   The General Register Office for Scotland was already making a lot of use of computers in the processing of vital events and production of analysis and statistics. These functions continue to be carried out on a mainframe computer located in Ladywell House, some 8 km to the west of New Register House, where the statutory registers and indexes are stored. When the project was being planned, a strategic decision was taken that a stand-alone mini-computer would be procured and be installed in New Register House, rather than the mainframe being upgraded and the system run 'remotely' from Ladywell House.

3.3.2   In view of the expected value of the procurement, Government regulations laid down that full 'open procurement' procedures had to be followed. Broadly speaking, this involved advertising the

project in the official journal of the European Communities, thereby inviting interested suppliers to obtain a copy of the Operational Requirement. Suppliers then submitted proposals, of which four were shortlisted. Following detailed discussions between GRO(S) and shortlisted suppliers, Invitations to Tender were issued and the contract subsequently awarded to the successful bidder. This is a time-consuming process, taking about a year from the drawing up of an Operational Requirement to the final award of contract. The process is seen to be fair, however, and should ensure that the delivered system represents best value-for-money for the purchaser, while giving all suppliers the opportunity of competing for the contract.

3.3.3 The mandatory requirements of the system as envisaged when the Operational Requirement was drawn up were briefly as follows:

i.   Able to read magnetic tape in specified formats, these containing the converted raw indexes

ii.   Capable of supporting up to 45 terminals operating concurrently

iii. Capable of supporting printers and terminals located in remote sites

iv.   Direct access storage of sufficient size for the operating system, system software, security files, programming and housekeeping, plus data files

v.   Capable of handling the specified volume of transactions; in this case, an estimated 600 transactions per hour at peak times by the end of year 5

vi.   Response times: between the operator hitting the 'enter' key and the required information appearing on the screen, not exceeding 5 seconds for 95% of all transactions

vii. Software: operating system capable of supporting concurrent transaction processing, batch processing and on-line program development; disk editing routines, direct data entry routines, facilities for transaction logging, facilities to secure against unauthorised access; data management routines for sorting, merging, dumping, printing, reorganisation and copying; file archival and restore facilities; application development software

viii. Capable of meeting agreed serviceability and reliability requirements

ix.   Physical specifications of screen and keyboard devices

x.   Printing requirements

xi.   Data archiving requirements

xii. Maintenance provisions

3.3.4 A great deal of time and effort was spent in ensuring that the sizing requirements of the system were understood and fully met by the proposed solutions of suppliers. Fast response times to user enquiries are essential if any system such as this is to achieve user acceptance.

3.3.5 The initial hardware configuration was installed in New Register House in December 1987. Since then, the system has been expanded on a phased basis to cope with increasing volumes of data to be stored and corresponding increase in on-line access to the data. Experience has been that the demand for terminals has been greater than originally forecast, with users being very keen to have ready access to a terminal wherever possible, rather than searching indexes using the paper volumes. Staff also benefit from not having to leave their workplace to consult the indexes.

3.4   Design and Development of an On-line Index Retrieval System

3.4.1 To recap, the paper-based indexes were subdivided into volumes by event types, year and sex; within these divisions entries were listed alphabetically by surname and forename, together with a Registration District (RD) name and entry number for each event. Certain other information was included to help distinguish entries with common names. A typical user would know the event type, the sex and approximate year of event, thus, narrowing down the search to one volume for each possible year. The user also had to know the surname of the entry being search for.

3.4.2 It was first of all decided that the data would be held in separate files for each year, event type and sex (that is, six files for each year); projecting ahead ten years from date of installation, this gave a total of 858 data files on the system (6 times 143 years). There were certain special characteristics of the data which allowed disc space requirements to be reduced by avoiding duplication. In particular, there was a great deal of duplication of names which would be held on file. Of the estimated 35 million records to be stored and accessed, many thousands of

records would have the same surname, Smith, for example. The system was therefore designed so so that the files of births, deaths and marriages hold a key number that points to the desired surname in a surname file, rather than holding the surname in each data record.

3.4.3   As many users would be unsure of the spelling of the surname they wished to access, a sound-alike search facility was also provided. The sound-alike code is stored with the corresponding surname in the surname file. By placing the surname and sound-alike code on a separate file, some extra processing is added to each transaction; this has a negligible effect on response times but saves considerably on disc storage requirements.

3.4.4   Once the raw index data has been converted and validated, it is passed to computer operations section to be loaded onto the computer database. It is at this stage that the surname and sound-alike codes are generated. If the surname is not already on file, a new code is created. As the historical record conversion progresses, however, fewer 'new' names are discovered. At present there are around 150,000 names on the surname file; this figure is well above what had been estimated before the project got under way.

3.4.5   Screen and dialogue design had to take account of the fact that many users of the system would have little or no previous experience of using a computer terminal. A high degree of user friendliness, with clear screen prompts and a simple, step-by-step approach, was essential. Two different on-line access applications were eventually developed; one for use by members of the public, and another for members of staff. Staff quickly adapted to the new skills required and did not need to use the easy-to-follow (but slower) method of accessing the data base.

3.4.6   Special keyboards were designed with the help of the equipment supplier to assist the inexperienced customer. Essentially, these were standard keyboards with most of the unused function keys physically jammed to prevent unintentional damage, and with special keys such as the ENTER and PAGE DOWN keys colour-coded for ease of reference. These keyboards were 'doctored' by our own staff using washers and sticky labels! They have proved very effective and were a much cheaper alternative to having the supplier provide 'customised' keyboards.

3.4.7   To access the database, the searcher will enter, via the keyboard, the type of event; year; sex; surname; and, optionally, first initial of first forename. All entries which match these criteria will be displayed on the screen; if the searcher finds the entry which he or she is looking for, the RD number and entry number can be noted down and a requisition slip completed. This slip can then be passed to a member of staff who will locate the register required. If the entry has not been found, the searcher is presented with a menu offering a number of options:

        - end the enquiry
        - re-display the search parameters to allow changes
        - search the year before the period already searched
        - search the year following the period already searched
        - display all records with similar sounding surnames
        - display all records with surname already searched but without specifying a first initial.

3.4.8   To allow corrections and additions to be applied to the database, certain members of staff, under password control, have authority to update files. For the most part, however, staff and members of the public are restricted to just viewing records. Full back-up copies of the database are taken regularly onto magnetic tape, for storage at an off-site location. To date the system has proved very robust, but a complete paper copy of the index is retained in the vicinity of the main searching area in New Register House for use in the event of serious system failure.

3.5   Accommodation

3.5.1   As mentioned previously, it was decided at the outset that a stand-alone mini-computer (an IBM System 38) should be procured and installed at New Register House to run the system. This meant that considerable upgrading of accommodation had to be carried out before the hardware could be installed. New Register House was opened in 1861 as a purpose-built building to hold the registers of births, deaths and marriages and certain other genealogical and population records. Not surprisingly, the metre-thick stone walls did not lend themselves readily to the installation of trunking to run computer cables throughout the building.

3.5.2   A room to house the computer was fashioned from an old store room in the basement of the building; facilities such as a dedicated power supply, false floor and air conditioning unit were

provided. Trunking was run throughout the building to allow for future expansion. Furniture suitable for use with Visual Display Units had also to be acquired for staff and customers who would be using the system. The introduction of computer technology into the searching process acted as a catalyst in getting GRO(S) to examine critically the use of space in New Register House. As a result, the reception and search areas in the building have subsequently undergone major refurbishment.

3.5.3 A separate section on accommodation matters is included here as a marker to any other registration Authorities who may be considering a similar project, and whose records may be currently stored in an older building. Accommodation requirements of such a system are a major item in their own right and require much careful planning at an early stage of the project, as well as careful monitoring of progress once development is under way.

4.   SUMMARY

4.1   Project Implementation

4.1.1 The final acceptance trials for the computer equipment were completed in March 1988. Staff have been using computer terminals to access the statutory indexes from that time. Staff have found the application easy to operate, and agree that there are considerable time savings over the old method of locating and looking-up bound volumes.

4.1.2 Members of the public started to use the computer-based index in March 1989. The application designed for use by customers makes use of easy-to-follow screen prompts; keyboards are 'customised' to include colour-coded keys and have superfluous function keys jammed and 'blanked-off' to prevent problems caused by unintentional pressing. Our experience has been that visitors to New Register House have adapted very quickly to the new technology. An initial introduction is given by the search room supervisor, and written guidance is available at each terminal for reference. Customers have then been able to use the computer system without further assistance from staff.

4.2   Problems

4.2.1 No major unforeseen problems have arisen during the course of the project. In chapter 3.2 of this paper, mention was made of the disappointing results achieved in data conversion using optical character recognition techniques. This was overcome by increasing resources allocated to the traditional key/verify method of data conversion. It did involve asking the financing authority for additional monies at an early stage of the project, however, something which one would obviously wish to avoid if at all possible.

4.2.2 The accommodation problems when introducing a system such as this will vary according to local circumstances. Points which have to be borne in mind include:

- Will a special purpose computer room be required, with raised floor, air conditioning, special power supply, etc.?
- Cabling requirements: location of terminals for use of staff and public searchers. In our case, special purpose trunking for both power and data cables was run throughout the building to give maximum flexibility.
- Bear in mind that until data conversion is completed, paper volumes and computer terminals must exist side by side.

4.3   Advice

4.3.1 In offering advice to other registration authorities interested in setting up such a project, we are aware that different procedures for civil registration exist in different countries. Solutions must therefore be designed to suit the particular needs of each authority. Nonetheless, there are a number of general principles which should apply to any system of computerised indexes.

4.3.2 Project Control: for a project this size there is merit in breaking it down into a series of sub-projects for planning, resource monitoring and control purposes. In our case, we chose to set up four sub-projects.

4.3.3 Volumes: determine carefully the volumes of paper index records you plan to convert, and decide on a realistic time-scale for data conversion.

4.3.4 <u>Accuracy</u>: devise a set of validity-checking programs which can ensure as far as possible that accurate data are held on the computer data-base. Users would soon lose faith in the system if entries were incorrect or frequently missing.

4.3.5 <u>Sizing</u>: assess carefully the access demands made on current paper indexes and determine likely growth patterns. This will be vital in ensuring that the computer hardware selected is sufficiently powerful to provide fast response times to queries.

4.3.6 <u>Ease of use</u>: bear in mind that the retrieval software must be sufficiently easy to use for someone who may have no previous experience of screen and keyboard devices.

4.3.7 <u>Communication</u>: where a project such as this involves the introduction of information technology into an area for the first time, staff must be kept fully informed on developments at all stages. A first-class technical solution could be of limited use if staff are not involved in setting up the application and committed to its successful implementation.

4.3.8 <u>Ergonomic factors</u>: again, where computer terminals may be being introduced for the first time, bear in mind the ergonomic factors associated with working with Visual Display Units, such as proper furniture, lighting requirements and cable management. Allowance may have to be made for these items in project costings.

5. FUTURE PLANS

5.1 General

5.1.1 The project to computerise the indexes to the statutory registers was conceived and implemented to allow the General Register Office for Scotland to continue to fulfill its statutory obligations in the face of increasing demand. The project is an important first step in the process of the registration service moving to a system of electronic storage and retrieval of records. This will give our customers a faster and more widely available service. At present, should a member of the public wish to obtain a copy of a birth, death or marriage certificate, he or she has three options:
    i.   call or write to the local registration office where the event was registered;
    ii.   call in person at New Register House in Edinburgh;
    iii. write to New Register House enclosing the necessary details.

5.1.2 The long term strategy of GRO (Scotland) is to move to a situation whereby members of the public can easily obtain a birth, death or marriage extract within 24 hours. As previously stated, the computerised index is seen as a first step in this process; two other major projects are envisaged to help us meet our objectives.

5.2 Computers in Local Registration Offices

A pilot study is currently underway in one local registration office, whereby registration details are collected via a screen and keyboard, rather than the traditional pen and paper method. This enables copy certificates and associated documentation to be produced automatically. It will also be possible in the longer term to give a local office on-line access to the central computer index in New Register House, Edinburgh.

5.3 The Electronic Register

This potential application is currently only at discussion stage. The concept is that the entire register of births, deaths and marriages for Scotland might be held on a central data-base which is accessible through terminals, using the computer index to locate the appropriate record. Such a system would have a number of significant advantages over the current book-based system, including faster searching, easier production of extracts, ability to locate terminals away from New Register House, and improved conservation of original historical records. There is much detailed preparation and planning to be done, however, before work could start on such a project.

5.4 Conclusion

As things stand at present, the GRO (Scotland) are satisfied that the decision to press ahead with computerisation of the statutory indexes has been fully justified. Productivity and customer service have been improved and the benefits of having a computerised index will continue to accrue long into the future.

8

PUBLICATIONS OF THE IIVRS TECHNICAL PAPERS