Developing a system of indexing surnames in the Home Office

John L. Rush

Formerly one of the deputies to the Chief Registrar of the Home Office

In The Indexer of April 1976, in his paper about Victoria's shipping records, Douglas Bishop mentioned that he found that the possible alternative ways of spelling the same surname was a common problem. If this is a problem in indexing archives, it becomes an even greater one in nominal indexes which are maintained for retrieval of current information, especially where an index is continually expanding and where many of the enquiries are made by telephone or are based on telephoned information. In the last case, of course, the risk of mis-hearing is added to the tendency of some enquirers to have preconceived ideas on how the name is or should be spelled.

In one of the nominal indexes kept in the Department we found that there was a problem of duplicated records—duplications which probably arose at different times—and a consequent risk that cases might be considered without all the facts being available. It is a large index kept in loose-leaf books and reference is made to its various sections about 600 times a day. It was then divided into eight alphabetical sections, and with experience, staff on each of the sections acquired a knowledge of the possible permutations of spelling they were likely to meet. However, new staff, or staff transferred from another section of the index, needed time to learn the new section. There was, in fact, a marked reluctance to move between sections and I only realized why when we began to examine the problems of similar names. We decided, therefore, that an attempt had to be made to incorporate in the index the knowledge of similar spellings possessed by the most experienced registrars. Little had been done elsewhere to produce the system we required—we found that in similar indexes other Departments also relied as we did on staff expertise. In fact, the only system which would meet our need was a phonetic system (the Remington 'Soundex'), which was designed for computer use. It was actually in use in one section of the Home Office but it was thought too complicated for a manual index. It had the added disadvantage that the filing system used by us depends on the initial letter of the surname, which would have caused confusion over names like Cain/Kane, Cavanagh and Phazackerly. We did hope that the computer would be able to produce a print-out from the phonetic names to the normal spellings. Had it been programmed to do so, much of the labour involved in compiling the index would have been saved.

We only learned of this system, however, when our own work was under way, and the fact that the 'Soundex' system used a 'master name' system (which we had independently developed) gave us confidence that our proposals would work.

We therefore decided that whatever was done had to be done independently. We considered whether it would be sufficient merely to produce lists of similar names, but decided that the necessity of continued reference would be irritating and time wasting. Once produced, though, these lists would be useful elsewhere purely as lists. As already mentioned, it had been decided that one name, or rather one spelling, would be selected as a 'master' name for the group, and all files indexed under it, however that name was spelled. The index is kept in loose-leaf form, one page per item and cross-reference sheets of a distinguishing colour were to be inserted in the 'correct' alphabetical sequence, referring the enquirer to the 'master' name where the individual could be distinguished by forenames, date of birth and other information.

Names like Cavanagh/Kavanagh, mentioned before, could not be grouped but were shown in each alphabetical section as a 'see also'. Similarly 'see also' was used for names which could not logically be grouped together but could be confused if spoken or written illegibly. Examples of this are Dagleish/Dalgleish, Paterson/Patinson or even Feasey/Pheasey/Veasey. These 'see also' references, together with all the 'grouped' names, are shown on a coloured 'master' names index sheet which precedes the individual file index sheets for the names in that group. Reference from the 'group' to the 'see also' section is, therefore, readily available. Some rules had to be made to reduce the number of names in the list and thus the volume of references and cross-references in the final index. Names with a final silent 'e'—Browne and Locke are obvious examples—were excluded and regarded as being covered by the shorter versions Brown and Lock. So were plural names, but names like Drake or Jones having no shorter version were included where necessary. Double barrelled names like Leveson-Gower or Nelson-Mallallieu are indexed under the last name. With names beginning Mac, Mc and M', the common practice of putting all three together is followed and this section is treated as a separate letter. In the list itself the prefix is omitted except for master names. This, we think, makes for
easy reference—it certainly reduces the typing load a little. In the index sheets, of course, the full name has to be used.

The longest and most arduous part of the work was compiling the list of names itself. It could only be done at times when full staff were available. It was delayed further because during the last few years much of the Home Office moved to a brand new building. Although not moving, we were responsible for planning and installing an electronic file storage system and this obviously carried a high priority. Since the staff engaged on the ‘master’ name work were as far as possible selected for length of experience and flexibility of outlook, they were obvious choices when other new work had to be done—either on the new scheme or keeping the system going with the staff left behind. I was amazed to find that with interruptions it has taken over three years to produce the draft index. This has not, of course, all been delay. Some sections were brought, completely or partially, over to the similar names system fairly early in the process, and the experience of new staff with these and with the old ‘strict alphabetical’ indexes was an encouragement to continue.

Although compiling the initial lists was undoubtedly the most difficult part, checking the names and putting them into a form which was compact enough for reference and for typing was equally laborious. Figure 1 shows a sample page of the final draft. This page contains one of the longer ‘master’ names—Jocelyn—with 26 cross-references. The largest number of variants we have noticed is Maconachie, with 66 variations.

We do not expect this list to be exhaustive. It is based on the knowledge of a number of experienced people but additions are still being made and will, no doubt, continue to be needed. The list contains mainly British names but foreign and Commonwealth names have been included as far as they are necessary in practice. It is in this area we can expect the greatest expansion.

As the index in which this system is being introduced is also the master index for several smaller indexes outside London, we have thought it necessary for the system to be fully set out and a copy of it will be included ultimately in the ‘Notes for Guidance’ which are given to the staff.

---

**INSPEC’S new Director**

The Institution of Electrical Engineers has announced that T. M. Aitchison, the Deputy Director of INSPEC, has been appointed Director, replacing D. H. Barlow who has resigned. Tom Aitchison has been with the Institution since the formation of INSPEC in 1967 and has been Deputy Director since 1971. The new Deputy Director is P. Clague, the former General Manager of INSPEC’s Database Operations, another foundation member of INSPEC.

---

**Indexing the world’s news**

The French newspaper *Le monde*, in collaboration with the publishing house F. P. Lobies, is recommencing publication of its yearly analytical indexes. To the eight volumes already published—those for 1944/5, 1946-49, 1965-67—will be added in 1980 indexes for 1968, 1950 and 1958, and the rest at the rate of three volumes a year until currency is reached.