(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
6 June 2002 (06.06.2002)

PCT

(10) International Publication Number
WO 02/45396 A2

(51) International Patent Classification[7]: H04M 3/42

(21) International Application Number: PCT/GB01/05224

(22) International Filing Date:
27 November 2001 (27.11.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
0028940.5 28 November 2000 (28.11.2000) GB

(71) Applicants and
(72) Inventors: **CRABBE, Anthony** [GB/GB]; 96 Dale Road, Matlock, Derbyshire DE4 3LU (GB). **MOLYNEUX, Pamela** [GB/GB]; 96 Dale Road, Matlock, Derbyshire DE4 3LU (GB). **MOLYNEUX, Hugh** [GB/GB]; 96 Dale Road, Matlock, Derbyshire, Derbyshire DE43 LU (GB). **CRABBE, Peter** [GB/GB]; 17 Harold Court, Reginald Street, Derby, Derbyshire DE23 8FR (GB).

(81) Designated States *(national)*: AU, BR, CA, CN, ID, IN, JP, MX, PH, PL, RU, US.

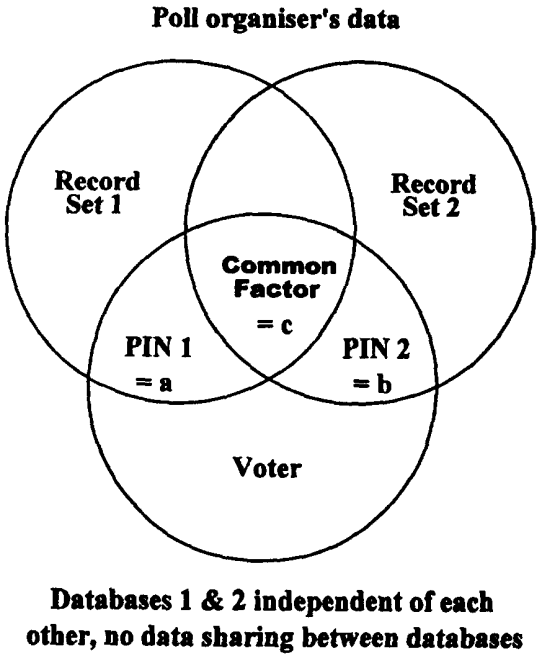
(84) Designated States *(regional)*: European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR).

Declaration under Rule 4.17:
— *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii)) for all designations*

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: SECURE TELEPHONE POLLING



(57) Abstract: Many automated systems already exist that allow voters to review options and cast votes using normal telephones. A secure telephone polling system enables users to vote anonymously from poll organisers and eavesdroppers, whilst still allowing authorised investigators to check that only legitimate voters have participated in the poll. The method by which a voter preserves his anonymity is to submit to the organisers, only parts of the symbol sequences that comprise his existing personal identification codes, such as his National Insurance number. A poll organiser requires two or more trusted third parties to provide only the same partial codes related to a single personal detail, such as a home postal code. When the organiser successfully matches two or more code entries to a single personal detail, this gives a very high probability that the person entering the codes is the same person fully recorded on the third party's original databases.

WO 02/45396 A2

SECURE TELEPHONE POLLING

BACKGROUND

In general, the organisation of polls may take three forms. They solicit votes in order to:

a) Confer a mandate upon selected representatives of the participating electorate.
b) Confer an honour upon selected individuals.
c) Register opinion about selected issues.

The organisation of a mandatory poll needs to protect both the privacy of individual voters and the poll against fraudulent or malicious individuals. The present invention is concerned primarily with enabling private polls of the type a) above to be conducted by telephone calls, but may also have application in the cases of b) and c). Principles of the present invention may also be applied more generally in cases where an individual is required to identify themselves through entry of Personal Identification Numbers (PINs).

1. Prior Art

There are numerous systems designed to handle telephone polls of types b) and c) above. In respect of a) above, patent searching reveals that existing solutions for validating a caller's subscription to a telephone service fall into three main categories:

1. Those that use telephone peripherals, such as a telephone card reader, which requires the caller to identify himself by use of a card issued by the service provider (US5412727, US4995081, WO96/02044)
2. Those that build recognition hardware into the telephone system, e.g. an identity chip on the caller's home telephone set, or a system that recognises the telephone set from which a call is being made (WO97/04602A2, US5838774, JP9081821A, JP8137969A, JP8044919A, WO99/26396)
3. Those that request the caller to enter PINs on the telephone keypad and match those PINs with pre-stored data about the caller (WO97/46031A1, US5689247, US5528670, US5311594, US3644675). The present system falls under this category.

Pilot tests for political telephone voting have been run in the USA and Canada in the 1970's and 80's.[1] The systems used there still relied upon elements of non-telephonic activity, such as postal correspondence. A feature of the present invention is that it can, in principle, all be operated by telephone, using a single call session to authenticate the caller and thereafter, to conduct a given poll. This feature is made possible by requiring the caller only to confirm personal data that is already pre-stored in existing public or commercial records.

To safeguard both the voter's privacy and anonymity, the present invention requires the caller to enter only fragments of the identifiers held in the pre-stored records, for example, the first six digits of an eight digit PIN. This use of identifier fragments allows the poll organiser a very high degree of certainty that a caller could not accidentally enter data matching their personal identifiers held in the pre-stored records. The use of identifier fragments entered on a telephone keypad also means that the caller need never give their name, address, or full PIN, which prevents anyone obtaining the poll organiser's records discovering the actual identity of the caller. Yet, the same data is sufficient to pinpoint a valid caller's postal area. The means by which the present invention achieves these said features, are now described by a combination of working principles and embodiments.

DESCRIPTION

1. Voter registration

Voter registration is the preliminary step to that of voting. In the preferred embodiment of the present invention, registration could be included as part of each voting event, or in a second embodiment, registration could be a one-off event that registered a caller as a voter for subsequent telephone voting in polls arranged by the same organisers.

2. Data Sources

For secret balloting, data about the prospective voter should be obtained from at least two, or more, independent public record sets, or databases. For open balloting, such as raising a petition by telephone, it may be acceptable to register a signatory with data drawn from just one database. If the signatory voluntarily provides their personal telephone number, then independent monitors can validate the petition's authenticity by calling back a given sample of signatories.

3. Secret balloting and data protection

For secret balloting, the following requirements are essential:

1. Only the voter knows what choice he or she has made.
2. The voter's personal details are neither shared nor disclosed in a manner contrary to data protection laws

The present invention addresses these secret ballot requirements in the following ways:

3. The voter never supplies their name, or address.
4. The voter only supplies a fragment of any given ID they have, for example, only the 6 digits of their 3 letter 6 digit UK National Insurance no. This prevents anyone with access to either the poll organiser's database, or the voter's telephone calls, from gaining enough information to consult other record sets in order to find out the voter's name.
5. The data fragmentation described in 2 above also means that the owners of public records can supply their records to a poll organiser, in a way that never breaches their Data Privacy obligations to individual citizens. The record owners may further secure their data from a poll organiser by "locking" the display mode of their computer database files, so that for instance, all file data is displayed in password format, ****. Proprietary software applications like *Microsoft Access* enable owners to set this type of data protection so that only users with the owner's password can change the file design.

4. Fraudulent voting and misuse of data

It is desirable to protect against the following polling abuses:

1. Fraudulent acquisition of another person's identity numbers.
2. Eavesdropping, such as telephone tapping.
3. Unauthorised use or distribution of individual records by the database holder.

The present invention does not allow either the poll organisers, or telephone eavesdroppers, to deduce the voter's name or address. However, the present system cannot prevent a) above if it is achieved by means such as mail interception, or disclosure by the voter, which is also a problem with other secure systems, such as credit cards and electoral registration. The present invention could make use of voice "signatures" if required. These voice entries could be recorded as WAV files, for instance. In the event of a fraud investigation the WAV files could be matched with recordings made by suspects.

5. Data Entry

The data entry and system responses for the preferred embodiment of the present invention are now described by example. The telephonic system linking the caller to the poll organiser via the telephone carrier is shown schematically in Figure 1. At the poll organiser's telephone exchange system, calls are relayed to a series of voice response interfaces, each linked to a personal computer, with each said computer being linked to a main server, in which the poll organisers keep their master database. The term "voice" may refer either to a human operator or a set of pre-recorded voice messages. The master database holds pre-stored personal identification data supplied by two record set holders who are independent of each other and do not share data. Callers are prompted to enter their details using either speech or the telephone keypad. Speech entries are recognised and processed either by a human operator or by voice recognition software installed on the controlling computer for the response interface. The communication between the response interfaces and computers would be managed by existing software, such as British Telecommunication's *Meridian* application, running on personal or main frame computers, linked either to an automated, or operator controlled telephone exchange. The sequence of registration procedures is then shown schematically in Figure 2.

6. Data Matching

In the following example, the poll organiser asks the caller to enter at least two individual ID numbers, a and c, where a is an element of a personal data set $p1$, stored in Database 1 and c is an element of a personal data set $p2$, stored in Database 2. Database 1 is owned by the Department of Social Security, an organisation which does not share any of its record data with the National Health Service, the owners of Database 2. Nor do the said owners share information through any intermediary such as the said poll organiser, because the said owners only supply the said poll organiser with fragments of the said sets $p1$ and $p2$.

Caller entry	Database 1
a = 1st 6 digits of National Insurance No.	Data field 1 = a = 1st 6 digits of N.I. No. Data field 2 = b = Postcode
	Database 2
c = 1st 6 digits of Nat. Health Service No.	Data field 1 = c = 1st 6 digits of NHS No Data field 2 = d = Postcode

As shown in the Venn diagram in Figure 4, the said poll organiser can match the said caller's entries a and c by finding a common factor, the said caller's postcode, in the intersection of sets $p1$ and $p2$. The general principle illustrated here is that the caller's personal data set $\{p\}$, can only qualify for inclusion in the registry of valid voters $\{V\}$ held in the poll organiser's master database, if it satisfies the following general criterion:

$$\text{For all } \{p\} \equiv \{(a,b),(c,d)\},\ \{p\} \text{ is a member of } \{V\} \text{ if and only if } b = d \text{ and } a \neq c \neq b \quad (1)$$

In this example, the total number n, of 6 digit sequences taken from an NHS ID, can only be 1 million. So potentially, at least 40 of the 40 million UK electors $\{V\}$, share the same 6 digit sequence for either an NI or NHS no. By coincidence, there may also be another 40 electors sharing one of the million or so valid UK Postcodes, $\{P\}$.[2] However, the odds against finding at random in $\{V\}$, a pair of NI & NHS 6 digit sequences (a,c) that both correspond to the same post code $b = d$, are

$$1/[(V/n_1)/V]\ [(V/n_2)/V]\ [(V/P)/P] = 1/(40/4\text{x}10^7)(40/4\text{x}10^7)(40/10^6) = 2.5 \text{ x } 10^{18} \text{ to } 1 \quad (2)$$

Since there are 40 million pairs of NHS and NI that do satisfy equation (1) above, then a rogue caller entering two 6 digit numbers at random has the following odds of getting his or her entry registered:

$$1 \text{ in } (2.5 \times 10^{18})/(4 \times 10^{7}) = 1 \text{ in } 6.25 \times 10^{10} \quad (3)$$

The security set-up of the present invention is then, based on a statistical notion of certainty. On the one hand, the use of PIN fragments helps to disguise the voter's identity. On the other hand, the criteria for relating the said PIN fragments give a very high level of confidence that they identify the same voter - and that the high odds against the registration of rogue voters effectively prevents them from participating in telephone polls.

7. Prior preparation of record sets

To safeguard against freak duplications in originating databases, it is necessary to search the said databases for duplicate values before using them for registration purposes. Before use, these said databases are filtered by date of birth, to remove all individuals under the voting age. Finally, each PIN is stripped down to 6 digits by removing the unwanted letters or digits in the manner illustrated in Figure 4. The removal of letters from the required data entry has the benefit of making data entry by telephone much easier for the caller.

8. Embodiment of a database set-up

Figure 5 shows a database set-up for the above embodiment. Callers enter the first six digits of their NHS number and the six digits of their NI number. For the purposes of example, the second columns of the NHS and NI data records are shown "unhidden". However, in practice, both these columns would be displayed in password format, as illustrated in Figure 5, and the design of the tables be "locked" in that view by the owner's choice of a 20 digit security password. The two query tables could also be locked in the same way, which still allows the database user to view the necessary query data.

PIN numbers are automatically assigned to every caller as their data is entered on the table called "Caller" in this example. In practice, the PINs will comprise of much longer digit sequences than those shown in the example. Each data field in the "Caller" table is set to reject duplicate data entries, so that each record of each call in which the caller seeks registration is unique and any caller entering the same identification details more than once cannot be registered more than once. In the example shown in Figure 5, caller 1 has entered erroneous information for their NI number and Caller 5 for their NHS number. Only callers listed in the "Match Postcodes" query will have their PIN numbers validated for use in the next phase, that of voting.

9. The voting process

This is the second of two processes, wherein registered callers can cast their votes in a poll. Figure 6 schematically shows the processes in the preferred embodiment that enable callers who have successfully registered themselves to cast their votes in a subsequent telephone poll. The said callers are guided through a menu of options, from which they may then make a selection by keying in the item numbers on their telephone keypad. As for the registration database, data fields are set to reject duplicate voter details and thus to prevent the same caller voting more than once.

10. Data Matching

As shown in the example database, illustrated in Figure 7, the voting options are defined on a table "Options for Election 001 and all votes cast by registered callers are entered on a form linked to a table, "Votes for Election 001. A sub-programme embedded in the form matches the caller's entries with the register of users and will not open the choice box in event of mismatches, which in turn, prevents the call from being recorded as an entry on the master database. The votes cast can then be counted and correlated with specific geographical areas by matching the individual votes with postcodes, as in the kind of crosstab query illustrated.

11. Data Privacy and Security

Figure 2 shows that the information given by callers during the voting process does allow the poll organisers to correlate the following personal information about the caller:

(1^{st} 6 of 10 digits of NHS no.) + (Postcode) + (Option choice).

But once again, the present system does not allow the said the poll organisers, or eavesdroppers to deduce the voter's individual identity, nor their name and individual address. So the present system provides a very high level of guarantee that the caller is the person who is described by public identification systems, and the system also secures the voter's right to anonymity. Only someone with legal authority to search all the databases used in the present system could reverse the odds to match the data with a particular individual. However, data protection legislation may allow the said poll organisers to supply to third parties, trend details abstracted from the above information, such as votes cast by geographic region.

12. Voter's check on how their vote has been recorded

A further benefit of the present invention over traditional voting systems is that voters can, if they wish, call the poll organisers to verify how their vote was recorded in a given poll. This they may do by calling another service, which operates as shown in the schematic of Figure 8. The voter dials the service number for the poll they wish to check and logs on by entering their registration PIN. The computer interface then automatically uses that PIN to searches the data table "Voter Cross Check" illustrated in Figure 8, matches the PIN with the vote option number and the name of that option. The option name is then announced by voice to the caller, via the response interface system.

CLAIMS

What is claimed is:

1. A system for a service provider to authorise a correspondent to be a legitimate user of the service without recording the personal identity of the said correspondent, wherein:
   a) the said provider records, on a database, the identification codes of any potential correspondent in a one-to-one relationship with at least one of the said correspondent's personal details, such as his home postal code
   b) a said correspondent enters onto the said database, only parts of the full sequences of symbols that comprise his said identification codes, such as the first six digits of an eight digit sequence
   c) the said provider authorises a said correspondent as a service user on the condition that two or more of the said correspondent's partial identification code entries match the same said personal detail that is related to each said personal identification code on the said database
   d) the said provider offers services such as voting to said correspondents who have entered said partial sequences of identification codes that satisfy the said condition for matching the said individual records on the said database.

2. The method according to claim 1 whereby the said service provider may authorise the said correspondent as a legitimate user of the said service, on the condition that one or more of the said correspondent's identification code entries match with the identification codes that already exist in a said one-to-one correspondence with the said correspondent's personal details on the said database.

3. The apparatus of claim 1, wherein the said service provider records data about said correspondents on a computer, using existing software that can automate actions and responses to and from the computer, including those said actions necessary for maintaining a telecommunication dialogue with the said correspondent.

4. The apparatus of claim 1, wherein a said correspondent enters his said personal identification symbols onto a computer database from a location remote from the said computer, by using a computer peripheral device such as a keyboard or a telephone.

5. The apparatus of claim 1, wherein the said service provider may use telecommunication devices such as telephone handsets, to present a spoken or printed menu of choice options to a correspondent who is authorised on the said service provider's database by the matching of said identification codes with said personal details.

6. The apparatus of the preceding claims, where a said correspondent may use the said peripheral devices to enter his choices of said options presented by the said service provider onto the said service provider's computer database.

7. A method where the said service provider may use computer database software to relate a number of said correspondent choice selections to the said correspondents' personal details, in order to produce summary information lists, such as those relating all the entries of one particular choice to one particular postal district.

8. A method wherein the said service provider obtains the data about said users from third parties who do not share their complete data with any other parties, who provide the said service provider only with parts of the said data, such as six of eight symbols from the said user's personal identification codes, and who provide identification codes that are related only to a user's postal code, not to his name or address.

9. A method where the said service provider cannot learn the full identity of a said correspondent who has been authorised by the methods according to the preceding claims, but can only identify the said correspondent as an anonymous person who has overcome high statistical odds against entering at random, one or more said partial identification codes that correspond with personal details supplied by the said independent third part data owners.

10. The method according to claim 9, where the said odds against random symbol entries matching said personal details increase with the number of symbols comprising a said personal identification code and comprising a said personal detail, such as a postcode.

11. The method according to claim 9, where the said odds against random symbol entries matching said personal details increases with the number of said third party databases stored by the said provider and therefore, the number of said partial identification code sequences that must be entered by the said correspondent.

12. A method where an authorised investigative agency may take the said service provider's database records and relate them back to the records of the said third party data suppliers to establish with a statistical probability that a particular person was the correspondent who entered a particular choice onto the said service provider's database.

13. A method wherein the said service provider may record the choices made by a correspondent and relate the said choice to the said correspondent's data set on a said database, in order that the said correspondent can use again the service described in the preceding claims, to check which choices have been related to his data set on the said database.

14. The apparatus of claim 13, where the said service provider may use computer software to record user choices and enable any said user to check the said records by using computer peripheral devices as described in the preceding claims.

15. A system as claimed in any preceding claim for a user to correspond with a remote service provider and to choose service options, such as voting for a political candidate, without disclosing their full identity to the said service provider.

16. A telecommunication voting system substantially as herein described and illustrated in the accompanying figures and diagrams.

Figure 1

Management system for secure telephone polling

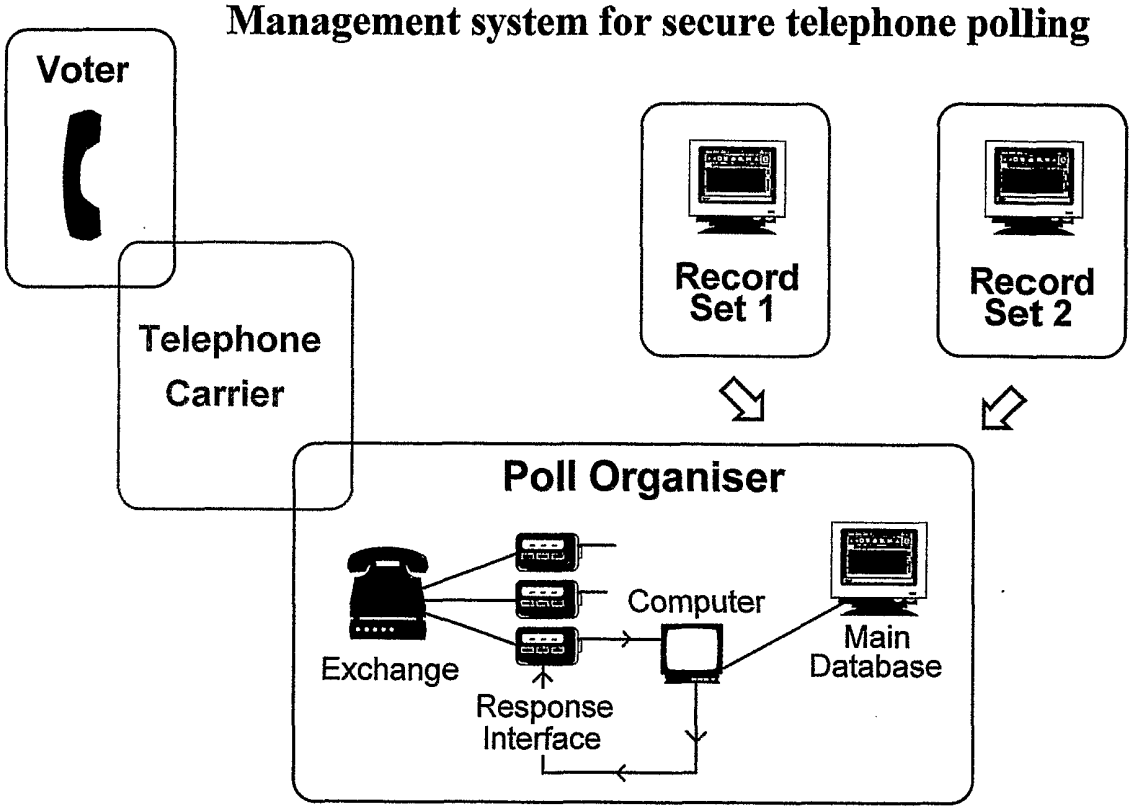


FIGURE 2.

REGISTRATION, Outline of data entry and system responses

Fill in "registration form" by dialling on keypad of touch-tone telephone.

	CALLER		SERVICE INTERFACE		DATABASE (ONLINE)
1	Dials service no.	→	Voice requests ID 1 no.		
2	Keys in ID 1 on phone pad	→		→	Logs ID 1, Date & Time of entry
3		←	Voice requests ID 2 no.		
4	Keys in ID 2 on phone pad	→		→	Logs ID 2, Date & Time of entry
5				↓	Verifies that ID 1 & 2 match the same postcode held in Databases 1 & 2
6A	Notes down PIN.	←	Voice read out of PIN.	←	Assigns PIN to correct matches and enters caller details onto register of users.
6B	Goes back to Step 2, or ends session.	←	Voice tells of data mismatch and asks caller to go back to Step 2, or end session.	←	Enters incorrect matches onto record of "invalid" users.

Figure 3

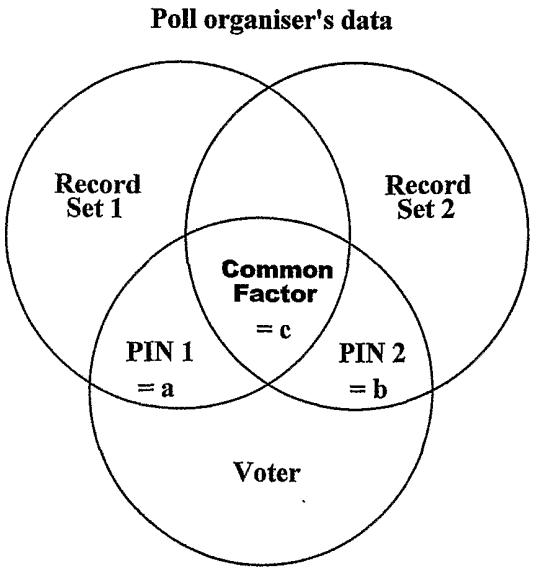


Figure 4

A method of extracting 6 digits from existing ID data

1. Illustration of different types of input masks used for inserting NI numbers onto databases. Masks within the database fields automatically split data forms like "GH315524D" and "0003783985" into the formats shown in the table below

ID	NI mask 1	NI mask 2	NHS mask
3	GH/315524/D	GH 315524 D	000378 3985
4	FE/460077/G	FE 460077 G	300789 0548
5	JU/471899/G	JU 471899 G	729787 1223
6	LP/862631/D	LP 862631 D	729511 5958
7	NB/493681/D	NB 493681 D	620238 9687
8	VB/129800/G	VB 129800 G	820033 1190
9	IU/526532/H	IU 526532 H	273600 9978
10	SX/233041/Y	SX 233041 Y	979991 6521
11	YT/517577/D	YT 517577 D	858483 1428
12	WR/546789/K	WR 546789 K	301301 9516

2. The records are copied from the *MS Access* tables above and pasted as unformatted text into an *MS Word* document

NI mask 1	**NHS mask**
GH/315524/D	000378 3985
FE/460077/G	300789 0548
JU/471899/G	729787 1223
LP/862631/D	729511 5958
NB/493681/D	620238 9687
VB/129800/G	820033 1190
IU/526532/H	273600 9978
SX/233041/Y	979991 6521
YT/517577/D	858483 1428
WR/546789/K	301301 9516

3. In *MS Word*, the unformatted text is converted into a table with columns split at "/" or "^"

NI Numbers Split

GH	315524	D
FE	460077	G
JU	471899	G
LP	862631	D
NB	493681	D
VB	129800	G
IU	526532	H
SX	233041	Y
YT	517577	D
WR	546789	K

NHS Numbers Split

000378	3985
300789	0548
729787	1223
729511	5958
620238	9687
820033	1190
273600	9978
979991	6521
858483	1428
301301	9516

4. Number values from the separated columns can then be copied and pasted back into "NI" & "NHS" *MS Access* databases used for registering voters.

Figure 5

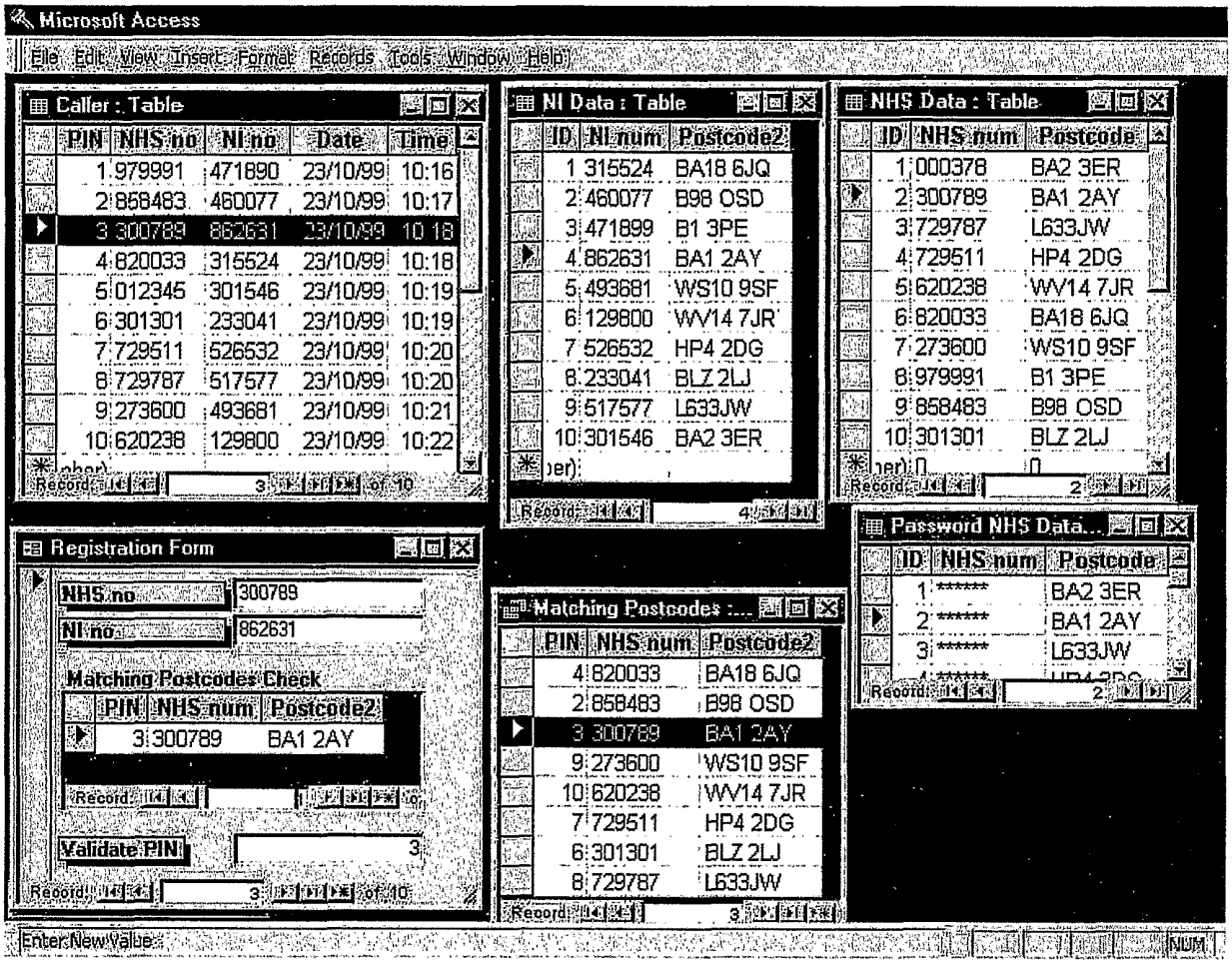


Figure 6

Option Selection

Fill in "option choice form" by speech or dialling on keypad of touch-tone telephone.

	CALLER		SERVICE INTERFACE		DATABASE (ONLINE)
1					Details of successfully registered callers are entered onto voting form
2	Notes options	←	Voice reads out menu of numbered options to registered caller	←	
3	Keys in menu choices	→		→	Logs choices made
4	Ends call	←	Voice notifies caller successful entry has been made	←	Notifies successful entry has been made

Figure 7

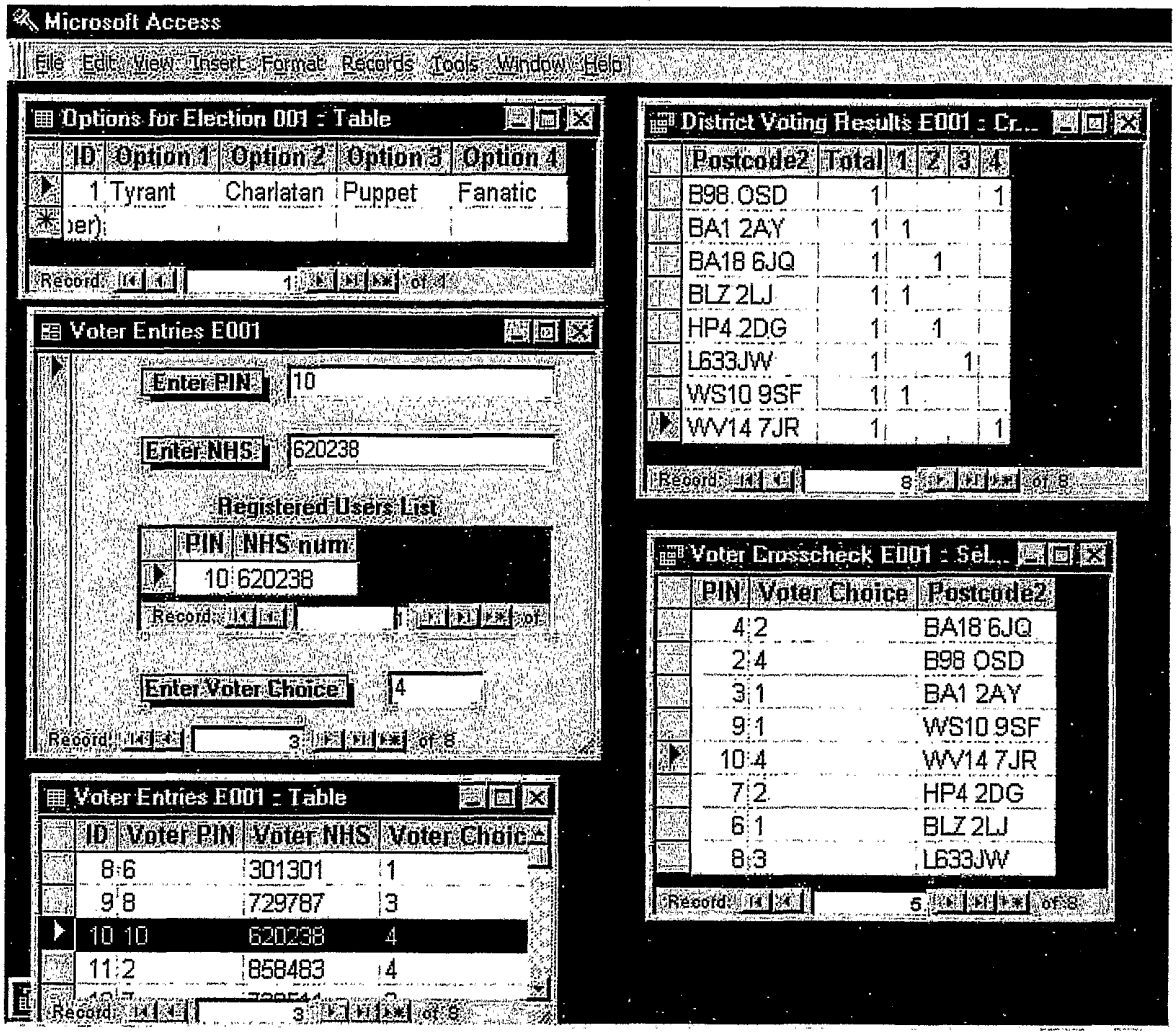


Figure 8

Voter confirmation system

	CALLER		SERVICE INTERFACE		DATABASE (ONLINE)
1	Dials service no.	→	Voice requests caller PIN.		
2	Keys in PIN on phone pad	→		→	Logs PIN.
3				↓	Searches Votes Cast records for match between PIN and choice number
4		←	Voice speaks the name chosen	←	Matches choice number with choice name