

La transformación del Banco de Datos del CEHME en una base de datos relacional

J. TOMÁS NOGALES FLORES
M. JOAQUINA FERNÁNDEZ SOLANA
Universidad de Extremadura

EL BANCO DE DATOS ACTUAL

La historia de este banco de datos se remonta al curso 1984-85, en el que, tras una serie de contactos entre el Dr. José Antonio Ferrer Benimeli, profesor de la Universidad de Zaragoza y especialista en el tema de la masonería, y el Dr. Antonio Rodríguez de las Heras, profesor de la Universidad de Extremadura, director del grupo de trabajo que por entonces llevaba el nombre de Seminario para la Investigación del Conflicto, equipo pionero en integrar la microinformática en el trabajo del historiador, se acuerda y se lleva a efecto la informatización del censo de masones de las logias de Valencia capital. Demostrada la gran utilidad de este banco de datos para los estudiosos del tema¹, por las facilidades que la informática aporta a la hora de poner en relación los miles de datos que contiene, en el curso 86-87 el Dr. Ferrer Benimeli propone la ampliación de su cobertura geográfica a todo el territorio nacional. El proyecto final obtiene la

1. El banco de datos es presentado en el *II Symposium de Metodología aplicada a la Historia de la Masonería Española. La Masonería en la España del siglo XIX*, organizado por la Universidad de Salamanca y celebrado en julio de 1985 en esta ciudad dentro de los Cursos Internacionales de Verano. Posteriormente, ya ampliado el banco de datos con la presencia de las logias andaluzas, efectuado un cambio de soporte hardware y software, e integrado en un proyecto financiado por la CAICYT (se verá más adelante), fue presentado en la ponencia *La campaña de informatización del censo de masones en Andalucía*, en el marco del *III Symposium de Metodología aplicada a la Historia de la Masonería Española*, organizado por el Centro de Estudios Históricos de la Masonería Española (Universidad de Zaragoza) y la Universidad de Córdoba, y celebrado en esta ciudad en junio de 1987.

financiación de la CAICYT, y se adquiere un microordenador que será dedicado exclusivamente a la gestión de este banco de datos².

Actualmente el banco de datos cuenta con aproximadamente cuarenta mil registros. Los datos van introduciéndose en la base de datos a medida que los investigadores de campo van recogiendo de las diversas fuentes repartidas por todo el país. Estos recogen los datos en fichas que posteriormente envían al SIC donde se someten a un proceso de adaptación al modelo informatizado, y se codifican los campos que lo requieren. Posteriormente se introducen en el ordenador y finalmente son sometidos a diversos procesos de ordenación y clasificación tendentes a eliminar en lo posible los errores producidos en el proceso, y sobre todo los inherentes a la propia fuente. Más adelante veremos en detalle por qué y cómo es ello posible.

Se ha considerado como unidad de estudio cada uno de los miembros de una logia masónica española. Por tanto, un registro del banco de datos contendrá los datos relativos a la estancia en una logia de un masón. Como consecuencia de ello, si un masón ha pasado por cuatro logias, existirán cuatro registros sobre él en el archivo³. En la figura 1 se reproduce la pantalla de introducción de datos, con datos ficticios, que ilustra sobre la estructura del registro y que puede servir también como referencia para seguir los comentarios que realizaremos a cada uno de los campos en que aquél está estructurado.

En esta ficha, cada dato relativo al masón, o lo que es lo mismo, cada campo de los que conforman el registro, aparece identificado por

2. Concretamente, un Amstrad PC 1512 con 20 Megabytes en disco duro, compatible con el estándar IBM PC/XT. El banco de datos estuvo soportado inicialmente en un ordenador Altos, bajo sistema operativo multiusuario Oasis, y manejado por medio del paquete de gestión de bases de datos relacionales Control. Con el cambio de soporte, el banco de datos hubo de ser rediseñado para adecuarlo al nuevo ordenador y al nuevo software de gestión de bases de datos, dBase III Plus, uno de los más potentes para los ordenadores basados en el sistema operativo MS/DOS (también llamados compatibles IBM PC/XT). Para evitar volver a introducir los mil quinientos registros que ya existían en el anterior soporte, correspondientes a las logias de Valencia, realizamos la conexión física de ambos ordenadores, en principio incompatibles (para ello debimos incluso fabricar un cable *ad hoc*, ya que no conseguimos encontrarlo en el mercado), y trasvasamos el banco de datos de uno a otro en un formato adecuado para la transmisión, tras la oportuna reconversión de los datos originales, y, por último, llevamos a cabo una nueva conversión de formato para conseguir que los datos recibidos se ajustaran a los requerimientos del nuevo programa de gestión de archivos.

3. Las razones para no optar por considerar como unidad de estudio al masón, sino en cuanto miembro de una logia, son varias. Entre ellas, la gran dificultad con que nos encontramos para afirmar que dos fichas de distintas logias corresponden al mismo masón: los errores de las fuentes, de muy diversos tipos (omisión de algunos datos o expresiones diferentes de uno mismo, sobre todo), nos impiden establecer con seguridad la identidad de un masón que haya pasado por diversas logias. Frecuentemente, sólo tras diversas ordenaciones y clasificaciones de los datos se hallan las claves para afirmar dicha identidad. Posteriormente se tratará este tema.

BANCO DE DATOS DE LA MASONERIA ESPAÑOLA.				CEHME - SIC
APell.: PEREZ GONZALEZ		NOMBRE: ANTONIO		
Simbol: COLON		Numero: 1		
SExo : M	EDad : 34	EstCiv: C	PROfes: COMERCIANTE	
FchNac: 1850/11/29				
Nacion: ESP	PrvNac: MA	MunNac: ANTEQUERA	LugNac: MA015	
ComAut: AN	ProVin: SE	MUNic.: SEVILLA	CIudad: SE091	
L.C.t.: L	LOgia : BETIS 12	OBedie: GODE		
AÑos : 1884/1886/1887/1888/1889/ / / / /				AñoIni: 1872
GRados: 03/18/18/30/33/ / / / /				
CarLog: TS /---/GT /OR1/VM / / / / /				
CauBaj: TRAS		FchBaj: 1889/08/23		
CarPol: DIPUTADO PROVINCIAL		ParPol: PSOE		
RElig.: CT				

Figura 1

un texto corto, seis caracteres como máximo⁴. Incluimos a continuación la relación de estos campos, con una descripción de su contenido y a veces un breve comentario; entre paréntesis figuran, tras la etiqueta asociada al campo, la longitud en caracteres de éste y una letra, A o N, que indica si se trata de un campo alfanumérico o numérico respectivamente.

APell. (24 A): Los dos apellidos del masón. En principio se expresan tal como aparecen en la fuente; si hay problemas de espacio, se ignoran, por este orden, la conjunción, las preposiciones, los artículos o el segundo elemento de apellidos compuestos; a este último caso se llega raramente.

NOMBRE (12A): El nombre del masón. Si es compuesto, del segundo sólo se expresa la inicial.

Simbol (15 A): El nombre simbólico del masón.

Numero (1 A): Si el simbólico ha de llevar ordinal (por repetirse en la logia), se expresará en este campo el numeral correspondiente.

SExo (1 A): Sexo, codificado con un carácter. Es interesante señalar, con respecto a este campo, que se ha optado por una codificación internacional: *M* para el sexo masculino (en inglés *male*, en francés *masculin*), *F* para el femenino (en inglés *female*, en francés *féminin*).

EDad (2 N): La edad del masón en el primer año de permanencia en la logia (éstos se relacionan en otro campo).

EstCiv (1 A): El estado civil, codificado con un carácter.

4. Las dos letras mayúsculas que aparecen siempre en este texto o etiqueta, tienen como función simplemente recordar al operador el nombre por el cual el programa de gestión de bases de datos que utilizamos reconoce al campo en cuestión. Estos nombres, de sólo dos caracteres para ahorrar pulsaciones, habrán de expresarse para hacer referencia a los campos a la hora de buscar, seleccionar, ordenar o clasificar los registros de la base de datos.

PRofes (15 A): Profesión. Mientras no contemos con una clasificación socio-profesional válida para el período de cobertura del banco de datos (finales del siglo XIX y principios del XX), las profesiones se recogen tal como aparecen en las fuentes⁵.

FchNac (10 A): Fecha de nacimiento del masón, ajustada al formato *aaaa/mm/dd*, donde *aaaa* es el año, *mm* el número del mes (01 para enero, 12 para diciembre) y *dd* el número del día (01 al 31). La conveniencia de utilizar esta notación para expresar una fecha se explica por el hecho de que basta una ordenación alfabética del campo en cuestión para conseguir su ordenación cronológica⁶. Si no se conoce el día, los dos últimos dígitos serán 00. Si tampoco se conoce el mes, los cinco últimos serán 00/00. Si se desconoce la fecha de nacimiento, el campo se deja en blanco⁷.

NAcion (3 A): Nacionalidad del masón, codificada con tres caracteres. No hemos podido utilizar los códigos normalizados existentes para la expresión de las nacionalidades actuales, por cuanto éstas no se corresponden en muchos casos con los de la época de cobertura de este banco de datos, ni en su denominación, ni en su demarcación. En vista de que la gran mayoría de los masones son de nacionalidad española, el programa de ordenador se encarga automáticamente, si el operador no expresa ningún dato para este campo, de incluir en él el código *ESP* correspondiente a la nacionalidad española.

PrvNac (2 A): Provincia de nacimiento, codificada con dos caracteres.

MunNac (20 A): El nombre del municipio de nacimiento. El contenido de estos dos campos aparece resumido en el siguiente.

LugNac (5 A): Lugar de nacimiento. Se expresará codificado con dos caracteres alfabéticos para la provincia y tres numéricos para el mu-

5. No parece fácil que se llegue a un acuerdo al respecto, si se ha de juzgar por los resultados de la mesa redonda habida al efecto en el marco del Symposium al que presentamos esta comunicación. En ella no se hizo sino poner de relieve las enormes dificultades de todo tipo que han de superarse para lograr una tal clasificación.

6. Debemos recordar aquí que, cuando se ordenan datos alfanuméricos, los números, tratados como simples caracteres, son anteriores a las letras y se ordenan entre ellos del 0 al 9. Por ello, habrá que expresar un cero delante del número de los meses o los días menores que 10, ya que, de lo contrario, la ordenación alfanumérica de fechas deja de convertirse en una ordenación cronológica (el mes 2 sería posterior al mes 12). Por otro lado, es interesante hacer constar que hemos preferido la utilización de un campo alfanumérico a uno de fecha: este tipo de campo está contemplado por la casi totalidad de los programas de gestión de bases de datos pero tienen algunos graves problemas, especialmente la falta de normalización (el formato interno de las fechas puede variar grandemente de unos programas a otros, originando problemas indeseables de falta de compatibilidad en los datos) y, en ocasiones, lo exiguo del período de tiempo contemplado (algunos paquetes de gestión de archivos no admiten la expresión de fechas anteriores a, por ejemplo, 1900). La utilización de un campo alfanumérico para las fechas siguiendo las directrices señaladas viene a solucionar estos problemas en la mayor parte de los casos.

7. Este campo está íntimamente relacionado con el campo edad (que, como dijimos, recoge la del masón en el primer año de afiliación a la logia); si se conoce alguno de estos datos puede deducirse el otro con una exactitud razonable (el error no será de más de un año arriba o abajo).

nicipio. Es este un buen ejemplo de campo codificado: para evitar por un lado el consumo excesivo de memoria en disco para un campo que toma un número de valores limitado (y por tanto susceptible de ser sometido a codificación), y por otro, los errores originados en la introducción de datos largos como son los nombres de algunos municipios y cuya expresión admite a veces varias formas (creando así problemas a futuras búsquedas de información en el banco de datos), este campo se ha codificado utilizando los códigos nacionales de los municipios, si bien con alguna modificación: se utilizan cinco caracteres, de los que los dos primeros, alfabéticos, representan la provincia (como en las matrículas de los automóviles, *BA* para Badajoz, *CC* para Cáceres, etc.) y los tres siguientes, numéricos, son el número oficial del municipio dentro de la provincia⁸. Es también un ejemplo interesante de la utilización de *máscaras* o *plantillas* a las que deben ajustarse los datos en el momento de su introducción en el banco de datos para que sean aceptados como válidos. Estas plantillas se definen en el momento de diseñar la ficha de pantalla. Al establecer la plantilla de este campo, se especificó que los dos primeros caracteres fueran letras (o letra y espacio) y los tres siguientes números.

ComAut (2 A): Comunidad autónoma actual en la que radica la logia a la que pertenece el masón⁹.

8. En la codificación oficial de los municipios españoles, la provincia está representada por dos dígitos: como ejemplo, el 06 para Badajoz y el 10 para Cáceres. Sin embargo, nosotros hemos preferido la codificación alfabética de la provincia, también con dos caracteres, atendiendo a una de las consideraciones que pueden hacerse en relación con las ventajas y características de la codificación: es deseable que el código asignado a un valor real sea un nemotécnico del mismo. Desde luego, los dos caracteres alfabéticos, que además responden a una codificación ampliamente extendida como es la de las matrículas de los automóviles, nos hacen más fácil la asociación con la provincia en cuestión que los dos caracteres numéricos. Y ello con la misma ocupación de espacio.

Llegados a este punto, merece un comentario la existencia simultánea de un campo codificado, *LugNac*, y otros dos, *PrvNac* y *MunNac*, vistos con anterioridad, que contienen la misma información sin codificar. Este aparente contrasentido tiene su origen en la historia, en la evolución de este banco de datos. En principio, constreñidos por la disponibilidad de espacio de memoria, hubo que hacer un uso intensivo de la codificación; esta exigencia se hizo menos imperiosa con el cambio de soporte informático al que nos hemos referido con anterioridad. Puesto que la codificación de localidades seguía siendo recomendable por los motivos comentados, se mantuvo el campo *LugNac*, pero como era interesante que en la ficha del masón (una copia de la pantalla de introducción de datos como la representada en la figura 1) apareciera este dato sin codificar, se añadieron otros dos campos, *PrvNac* y *MunNac*. Idealmente, si se tratase de una base de datos relacional, el operador no tendría que introducir más que el código del municipio: el programa se encargaría automáticamente de localizar el nombre del municipio en un archivo y de presentarlo en el lugar correspondiente de la ficha. Pero para que esta solución fuera viable, la base de datos habría de tener un volumen que compensara la existencia de un archivo que contuviese los nombres y códigos de los más de ocho mil municipios extremeños. Ahora es el momento de afrontar esta tarea, si bien no vamos a hacer mención de este *cuarto archivo* de la base de datos relacional para no complicar innecesariamente su descripción.

9. La existencia de este campo se debe a la división en campañas de la informatización del censo de la masonería, campañas que se han establecido para cubrir una o varias comunidades autónomas, teniendo en muchos casos como meta la publicación de libros o la elaboración de tesis doctorales. La necesidad de manejar la información (buscar, seleccionar, ordenar o clasificar datos) atendiendo a estas campañas hicieron aconsejable la inclusión de este campo. Los códigos de dos caracteres que hemos asignado a las comunidades autónomas son *AN* (Andalucía), *AR* (Aragón),

ProVin (2 A): Provincia en la que se halla la logia, codificada con dos caracteres.

MUnic. (20 A): Nombre del municipio donde radica la logia.

Cludad (5 A): El código de la localidad donde radica la entidad masónica. El sistema de codificación es el mismo utilizado para el campo *LugNac* (lugar de nacimiento). Este campo recoge codificada la misma información que los dos anteriores, *ProVin* y *MUnic*, como sucediera con los campos *PrvNac* y *MunNac* con respecto a *LugNac*. Los comentarios que realizamos a propósito de este último pueden aplicársele también a *Cludad*.

L.C.t. (1 A): Se expresará el código (normalmente la inicial) del tipo de entidad masónica (Logia, Capítulo, Triángulo...).

Logia (25 A): El nombre de la logia, y, si lo hay, su número. Si a lo largo de su existencia ha tenido varios, por cambiar de obediencia, se expresarán todos ellos, separados por barras.

OBedie (18 A): La obediencia de la logia. Si a lo largo del tiempo ha tenido varias obediencias se expresarán todas ellas separadas por barras. La obediencia está codificada con sus siglas (*GODE*, *GOE*, *GOI*, *GOLU*, etc.).

AÑos (49 A): Los años para los que existe constancia de que el masón ha estado afiliado a la logia. Van separados entre sí por una barra¹⁰.

AñoIni (4 A): Año de iniciación del masón.

GRados (29 A): El grado alcanzado por el masón. Se expresarán con dos caracteres (grados *01* al *33*), separados por barras, y estarán referidos, como en el caso de los cargos (campo *CarLog*, que figura a continuación), a los años de pertenencia a la logia (campo *AÑos*).

CarLog (39 A): Los cargos en la logia, codificados con tres caracteres y separados por barras, tenidos en cada uno de los años de permanencia en la logia. Para los años en los que no se tuviera ningún cargo se utilizará la convención «*---*»¹¹.

CauBAj (4 A): En su caso, causa de la baja del masón en la logia, codificada con cuatro caracteres.

AS (Asturias), *CB* (Cantabria), *CL* (Castilla y León), *CM* (Castilla-La Mancha), *CT* (Cataluña), *EU* (Euskadi), *EX* (Extremadura), *GA* (Galicia), *IB* (Islas Baleares), *IC* (Islas Canarias), *MA* (Madrid), *MU* (Murcia), *NA* (Navarra), *RI* (La Rioja), *VA* (Valencia).

10. Un cálculo simple nos muestra que hay espacio en los 49 caracteres del campo para contener diez años con sus nueve separadores. Los dos campos siguientes, grado y cargo en la logia, permiten asimismo introducir hasta diez datos con nueve separadores. Apenas una decena de casos han excedido estas previsiones iniciales, lo que, en un volumen de registros como el de este banco de datos, no supone una limitación apreciable.

11. Si un año se ha tenido más de un cargo, se repetirá el año las veces necesarias en el campo correspondiente, con el fin de respetar la correspondencia de los cargos con los años (lo dicho sirve también para el campo del grado, *GRados*, visto anteriormente).

FchBaj (10 A): Fecha de la baja.

CarPol (24 A): Cargos políticos que hubiera ejercido el masón durante los años de afiliación a esa logia concreta.

ParPol (4 A): Partido político al que, en su caso, hubiera estado afiliado durante ese período de tiempo.

RElig. (3 A): Religión profesada por el masón, si la hay y se conoce¹².

Merece un comentario la posibilidad que nos brinda el tratamiento informatizado de archivos para detectar la alteración, el *ruido*, que ha sufrido la información original en los diversos trasvases y conversiones a que ha sido sometida desde que se plasma en la fuente hasta que se introduce en el ordenador, ambos momentos incluidos. Entre aquellos trasvases y conversiones a que ha sido sometida desde que se plasma en la fuente hasta que se introduce en el ordenador, ambos momentos incluidos. Entre aquellos trasvases pueden señalarse, al menos, el paso de la información original a la fuente, de ésta a la ficha intermedia en que la recoge el investigador, y de ésta al ordenador. Cada paso va generalmente añadiendo errores, ruido. Una parte de esos errores serán de imposible detección (y es algo que debe reconocerse y aceptarse), pero otra, a veces muy grande, sí que puede detectarse y corregirse, siempre que exista un grado suficiente de redundancia en los datos.

El ordenador puede ayudarnos, aprovechando al máximo dicha redundancia, a detectar y corregir los errores introducidos por el inevitable ruido en cualquier parte del proceso. Incluso algunos de los errores, inexactitudes y ausencias en la información que nos proporciona la fuente, pueden ser detectados. Vamos a comentar brevemente uno de los múltiples sistemas que pueden utilizarse para conseguirlo, sobre el caso concreto de este banco de datos.

Como se ha dicho, la unidad de estudio, el registro, en este banco de datos es cada uno de los masones afiliados a una logia. En principio, para seguir la trayectoria de un masón, su paso por distintas logias, bastaría con conocer la relación de registros que contuvieran, en los campos correspondientes, los apellidos y el nombre del masón en cuestión (y, por ejemplo, el nombre simbólico, para aquellos raros casos en que coinciden nombre y apellidos en dos individuos diferentes). Pero en la práctica, esto es poco o nada factible, por cuanto en muchos casos un mismo masón aparece referenciado de diferentes maneras en cada una de las logias por las que ha pasado. Unas veces se expresa únicamente el primer apellido, otras los dos, otras sólo la inicial del nombre o uno de ellos si es un nombre compuesto, y muy frecuentemente cambia la expresión

12. Estos tres últimos campos, previstos en el registro desde un principio pero todavía con contenido en un débil porcentaje de casos, se irán completando más adelante y a partir de otras fuentes, complementarias de las que se están utilizando actualmente.

bien de unos, bien de otros. Y esto ciñéndonos a los campos citados. En estas condiciones es muy difícil por los medios tradicionales advertir que bajo diversas denominaciones se trata de un mismo individuo. Generalmente estas diferencias se han establecido en las propias fuentes, pero el ordenador nos permitirá tratar de la misma manera aquellos casos en los que tales diferencias procedan de los restantes trasvases que ha sufrido la información (de la fuente a las notas del investigador y de éstas al ordenador).

Un sistema elemental para advertir la presencia de estos casos, pero que da grandes resultados a pesar de su sencillez, consiste en la ordenación sucesiva de la información del banco de datos atendiendo a diversos campos. Una tarea impensable, por ardua, para ser realizada por los medios tradicionales, que sin embargo el ordenador realiza en pocos minutos. Una vez que se obtiene una lista ordenada alfabéticamente por apellidos y nombre de todos los registros del banco, y en la que figuren junto a estos campos algunos otros bien significativos como el simbólico, el lugar y/o la fecha de nacimiento, no hay más que echarle un vistazo par descubrir los casos en que el segundo apellido falte o se haya expresado de diferente manera puesto que todos estos casos aparecerán próximos entre sí en el listado. Para los casos en que sólo figura el primer apellido, habrá que buscar entre los que tienen además de éste un segundo : cuando encontremos una coincidencia en el nombre y el simbólico tendremos una gran probabilidad de que ambos registros correspondan al mismo individuo; si además coinciden lugar y año de nacimiento (u otros campos que se hayan incluido en el listado y puedan individualizar suficientemente), tendremos una razonable certeza de esa identidad.

Si el error, la diferencia en la expresión (o la ausencia de información), está en el nombre, el simbólico, u otro campo presente en el listado, este sistema nos permitirá detectarlo de la misma manera. El paso siguiente, aceptada la identidad de un masón, será la corrección del dato erróneo o la inclusión del dato que no aparecía expreso en alguno de los registros¹³.

Esta primera ordenación no nos permite detectar diferencias en la expresión de los datos de identificación de un individuo cuando tales

13. Un ejemplo sobre lo dicho: supongamos que en el listado ordenado por apellidos aparecen dos registros, muy próximos, correspondientes a unos tales «Maesa Oliva, José D» y a «Maeso, José D», ambos con simbólico «Mazzini» y profesión «empleado», dándose además la circunstancia de que coincide su fecha de nacimiento, aunque sólo se conoce el lugar de nacimiento y el estado civil de uno de ellos. Puede afirmarse con certeza (casi) absoluta que ambos registros corresponden al mismo individuo. De esta manera podemos completar el campo del lugar de nacimiento del registro en que no figuraba en este dato, y completar el segundo apellido, «Oliva», de aquel que sólo tenía el primero. Respecto del primer apellido, la diferencia en la expresión, «Maesa»/«Maeso», debe unificarse para mantener la identidad del masón. Puede escogerse «Maeso» por ser más plausible como apellido (al menos si atendemos a la frecuencia con que aparece en el banco de datos). Sin embargo, y ello es obvio, no deben completarse aquellos campos que pueden cambiar con el tiempo, como el del estado civil que conocemos para uno de los dos registros.

diferencias se hallen en el primer apellido (especialmente si está en sus primeras letras), puesto que van a aparecer muy separados en el listado. Se recurre entonces a una nueva ordenación, atendiendo ahora a otro campo. Este debe reunir dos requisitos: que figure en todos o al menos una gran mayoría de registros (con lo que excluiríamos aquellos que sólo se expresan esporádicamente, como la fecha de nacimiento), y que su contenido no se repita demasiado en los distintos registros (excluyendo así campos que no van a servir como criterios para personalizar o individualizar a un masón, como el sexo o el estado civil), requisitos que, como se ve, cumplían también los campos que se utilizaron para la primera ordenación, los apellidos y el nombre.

Un buen candidato para la segunda ordenación va a ser el campo del simbólico, dato que se conoce para casi todos los registros y que no suele repetirse demasiado. Así pues, los grupos que establece una clasificación por este campo son lo suficientemente pequeños como para que resulte cómodo buscar en ellos similitudes en los restantes campos que puedan orientarnos sobre la identidad de dos o más masones. Si además de la coincidencia de símbolo se da la de otros campos significativos, presentes en el listado, puede llegar a firmarse con una certeza razonable que determinadas diferencias menores en la expresión de apellidos y nombres (en el primer apellido, sobre todo, ya que es el que no queda cubierto en la primera ordenación) no obedecen a una identidad distinta del masón sino a un error o una omisión en algún punto de la trayectoria seguida por la información original ¹⁴.

Para cualquiera de estas dos ordenaciones (o más, si se estiman necesarias), la identificación del dato erróneo no suele ser en exceso problemática. Únicamente en los casos en que sólo se dispone de dos expresiones diferentes del mismo dato y ambas son igualmente aceptables ¹⁵ puede tenerse la duda de cuál considerar válida. Si hay tres o más, podremos considerar correcta la expresión más repetida.

Sin embargo, hay que hacer hincapié en la necesidad de que se dé esa certeza razonable de que nos encontramos en presencia del mismo individuo. A veces, la igualdad de nombre y apellidos, e incluso simbólicos, en dos registros diferentes no se explica por una misma identidad: los

14. Otro ejemplo, extraído también como el anterior de la experiencia real, aclara la utilidad de esta segunda ordenación. Dos registros que en la primera ordenación quedaban muy alejados son contiguos en esta segunda: corresponden a «De Olmedo López, Antonio» y «Olmedo López, Antonio»; puesto que el simbólico es el mismo en ambos casos, «Lope de Vega», aparecen próximos en el listado, con lo que es fácil advertir que se trata del mismo individuo pero con una expresión diferente del primer apellido.

15. Como ejemplo, dos registros correspondientes al mismo masón en el que se ha expresado de diferente manera uno de sus apellidos: «Giménez» y «Jiménez». En estos casos, para mantener la identidad del masón, si está probada, se impone unificar la expresión del apellido (por ejemplo como «Jiménez»), que puede posteriormente ser modificada si así se precisa en vista de los datos aportados por nuevos registros del mismo masón (esto es, si los nuevos registros introducidos hacen más frecuente la expresión del apellido como «Giménez»).

registros en cuestión pueden corresponder a un padre y un hijo (con el mismo nombre que el padre, caso frecuente, y el mismo segundo apellido, poco frecuente pero posible) que, por tradición, ha adoptado el mismo simbólico de aquél. El resto de los campos significativos, como ya se dijo, entre ellos el lugar y la fecha de nacimiento, la edad, la profesión o el año de iniciación, nos darán la clave (cuando se conocen para ambos registros) para asegurar que los dos (o más) registros corresponden al mismo masón. Si por el contrario existe alguna duda razonable de que esto sea así, es preferible por supuesto mantener los datos originales a alterarlos.

LA RECONVERSION DEL BANCO DE DATOS

¿Qué queremos cambiar de este banco de datos? Y, sobre todo ¿por qué queremos cambiarlo?

Si observamos la ficha o pantalla de introducción de datos que figura reproducida líneas atrás (fig. 1) y tenemos en cuenta que cada registro contiene los datos relativos a la afiliación de un masón a una logia, comprobaremos que algunos campos contienen una información que va a ser fija, invariable, para todos los registros que hagan referencia a un mismo masón (su nombre y apellidos, sexo, fecha y lugar de nacimiento o año de iniciación) o a una misma logia (su nombre o localidad en que radica), habiendo otros datos, eso sí, que podrían cambiar o de hecho cambiarían de un registro a otro, aun tratándose del mismo masón (su profesión, estado civil, edad, años en que ha estado afiliado, grados y cargos masónicos) o de la misma logia (su número de inscripción o su obediencia)¹⁶.

Esto es, si en el banco de datos hay diez registros sobre el mismo masón (porque estuvo en diez logias distintas), aquellos datos que hemos citado como invariables aparecerán diez veces en la base de datos. Más llamativo es el caso de los datos relativos a las logias: si hay mil masones en una logia, en la base de datos estarán repetidos mil veces los datos invariables de dicha logia (en los mil registros de los masones afiliados a ella). Y esto no sólo constituye un despilfarro de espacio en memoria de disco, sino que además, para el operador que introduce los datos por

16. Sin embargo estos datos se han considerado invariables en la solución que hemos dado hasta ahora al problema. Como vimos anteriormente, en la base de datos actual se expresan todas las diferentes obediencias (lo mismo se aplica a los números de inscripción) que haya podido tener un taller masónico, separadas entre sí por una barra. En el archivo de logias que propondremos, cada logia podría tener tantos registros como obediencias hubiera tenido. Es algo a estudiar con mayor detalle.

medio del teclado, supone un gasto innecesario de tiempo y un mayor número de pulsaciones.

La conversión del archivo actual de la masonería en una base de datos relacional viene a solucionar esta indeseable situación. Dicho de la forma más sencilla, una base de datos relacional es un conjunto de archivos que están relacionados entre sí, estableciéndose el vínculo entre dos registros de dos archivos diferentes cuando determinado campo (o combinación de campos) de un registro, tiene el mismo valor (contiene el mismo dato) que otro campo, también preestablecido, del otro registro.

En nuestro caso, contábamos con un archivo único cuyos registros contenían los datos relativos a la afiliación de un masón a una logia. Se trata ahora de conseguir una base de datos relacional que cuente, es nuestra propuesta, con tres archivos. El primero contendrá en sus registros los datos de afiliación de un masón a una logia concreta; este primer archivo no contendrá pues los datos que hemos considerado como invariables, salvo, únicamente, los que servirán para poner en relación a los registros de este archivo con otros registros de los dos archivos restantes. Estos otros contendrán los datos que hemos considerado como invariables, recogiendo los registros de uno de ellos los datos de cada logia y los del otro los relativos a cada masón.

El primer archivo, el de afiliación del masón a una logia concreta, contiene obviamente datos relativos al masón y a la logia. Acerca del masón tenemos los datos que hemos llamado variables (edad, estado civil y profesión, y el ordinal de su nombre simbólico en la logia) y tres campos con los datos invariables que definen (casi) inequívocamente al masón (nombre, apellidos y simbólico) y que por tanto serán los que se utilicen para relacionar este registro con el correspondiente al individuo en cuestión en el archivo que podríamos llamar biográfico y que contiene, además de estos tres datos, los que hemos denominado invariables (sexo, fecha y lugar de nacimiento o año de iniciación). El archivo de afiliación del masón a una logia contiene sólo un dato acerca de la logia: el campo con el nombre y número (o números) de inscripción de la logia, suficiente para identificar (casi) inequívocamente a la logia y por tanto para relacionar este registro con el correspondiente a la logia en cuestión en el archivo de entidades masónicas, que será el que contenga, además de este dato, todos los demás relativos a la logia, invariables en los términos definidos con anterioridad (tipo de entidad masónica, obediencia¹⁷ y localidad en que está ubicada). Este primer archivo de nuestra base de datos relacional se completa con los datos propios de la afiliación del

17. Conviene recordar aquí lo que hemos apuntado poco antes: la logia puede cambiar de obediencia pero en la solución que hemos adoptado hasta ahora se recogían en el mismo campo la lista completa de obediencias que hubiera tenido la logia, y, en relación con cada una de ellas y en el campo del nombre de la logia, los números de inscripción correspondientes.

masón a la logia, entre ellos los años de permanencia, grados alcanzados y cargos desempeñados en el taller masónico.

La figura 2 recoge, con un diseño afín al de la pantalla de introducción de datos reproducida en la figura 1, el esquema de un registro de este archivo, con los campos que lo formarían.

BANCO DE DATOS DE LA MASONERIA ESPAÑOLA.		CEHME - SIC
APell.: PEREZ GONZALEZ	Nombre: ANTONIO	
Simbol: COLON	Numero: 1	
EDad : 34	EstCiv: C	Profes: COMERCIANTE
LugRes: SE091		
LOgia : BETIS 12		
AÑos : 1884/1886/1887/1888/1889/ / / / /		
GRados: 03/18/18/30/33/ / / / /		
CarLog: TS /---/GT /OR1/VM / / / / /		
LogAnt: HISPALIS 8		
CauBaj: TRAS	FchBaj: 1889/08/23	
ARchiv: ARCH. HIST. SALAMANCA		

Figura 2

Como puede observarse, hay algunos campos nuevos que recogen datos que no aparecían en el diseño actual del banco de datos, y que proponemos para su inclusión por el interés que puedan tener. Son los siguientes:

LugRes (5 A): Lugar de residencia del masón, codificado de la misma forma que los campos *LugNac* y *Ciudad*, tratados anteriormente.

LogAnt (25 A): Logia a la que, en su caso, hubiera estado afiliado anteriormente el masón, antes de ingresar en la actual.

ARchiv (40 A): Archivos y/o fuentes de donde se ha extraído la información que figura en el registro. Sería conveniente estudiar una codificación adecuada para este campo.

En la figura 3 se representa un registro del segundo archivo de la base de datos relacional propuesta, el archivo que contendría los datos relativos a las entidades masónicas. El campo *LOgia* serviría de vínculo entre éste y el primer archivo.

BANCO DE DATOS DE LA MASONERIA ESPAÑOLA. ENTIDADES MASONICAS.		CEHME - SIC
LOgia : BETIS 12	LCTka : L	OBedie: GODE
ComAut: AN	ProVin: SE	MUnic.: SEVILLA
		Ciudad: SE091
FchFun: 1880/03/13	N.Fund: 11	N.Miem: 89
FchCes: 1898/12/17	CauCes:	MasFem: M

Figura 3

También en este archivo se aprecian algunos campos que no existían en el registro del banco de datos actual. Se relacionan a continuación.

FchFun (10 A): Fecha de fundación de la logia.

N.Fund (3 N): Número de miembros fundadores.

N.Miem (4 N): Número de miembros con que ha contado a lo largo de su existencia.

MasFem (1 A): Con un único carácter se indica, de forma codificada, si la logia es exclusivamente masculina o femenina, o bien de alguna forma mixta (esto es, si las mujeres son admitidas en pie de igualdad o sólo como miembros de adopción).

FchCes (10 A): Fecha de cese o cierre de la logia.

CauCes (4 A): Causa del cierre de la logia, codificada con cuatro caracteres.

La figura 4 recoge lo que podría ser un registro del tercero de los archivos de la base de datos relacional que proponemos. Sus campos contienen datos relativos al masón, tanto los invariables (entre ellos los apellidos, el nombre y el simbólico, cuya combinación sirve de enlace con el primer archivo, el de afiliación del masón a la logia) como algunos otros que aun pudiendo variar con el tiempo (cargo político, partido político o religión) es interesante recoger en este archivo biográfico en lugar de en el de afiliación, especialmente si se tiene en cuenta que estos datos se conocen sólo para un exiguo número de casos.

BANCO DE DATOS DE LA MASONERIA ESPAÑOLA. ARCHIVO BIOGRAFICO.			CEHME - SIC
APell.: PEREZ GONZALEZ		Nombre: ANTONIO	
SImbol: COLON			
SExo : M			
FchNac: 1850/11/29		FchDef: 1929/09/20	
NACion: ESP	PrvNac: MA	MunNac: ANTEQUERA	LugNac: MA015
AñoIni: 1872		GraMax: 33	
CarPol: DIPUTADO PROVINCIAL		ParPol: PSOE	
RElig.: CA			

Figura 4

También aquí aparecen algunos campos que no figuraban en el diseño actual del banco de datos de la masonería. Los siguientes:

FchDef (10 A): Fecha de defunción del masón.

GraMax (2 A): Grado máximo alcanzado en la masonería.

De lo que hasta ahora hemos venido exponiendo podría sacarse la errónea impresión de que, con el nuevo diseño del banco de datos de la masonería española como base de datos relacional, habremos de mantener

tres archivos, con sus tres pantallas de introducción y modificación de datos, en lugar de uno solo. Sin embargo, esto no es exactamente así. A los ojos del usuario puede seguir existiendo, si así interesa, una sola pantalla de introducción de datos. Podría ser la que se reproduce en la figura 5, o bien otra en la que se estableciera una separación o distinción más evidente entre los campos, según a cuál de los tres archivos pertenecen y cuáles tienen la función de establecer los enlaces entre ellos.

BANCO DE DATOS DE LA MASONERIA ESPAÑOLA.				CEHME - SIC
APell.: PEREZ GONZALEZ		NOMBRE: ANTONIO		
SImbol: COLON		NÚmero: 1		
SEXo : M	EDad : 34	EstCiv: C	PROfes: COMERCIANTE	
LugRes: SE091				
FchNac: 1850/11/29				
NACion: ESP	PrvNac: MA	MunNac: ANTEQUERA	LugNac: MA015	
FchDef: 1929/09/20				
LOGia : BETIS 12		LCTka : L	OBedie: GODE	
ComAut: AN	ProVin: SE	MUNic.: SEVILLA	CIudad: SE091	
FchFun: 1880/03/13	N.Fund: 11	N.Miem: 89	MasFem: M	
FchCes: 1898/12/17	CauCes:			
AÑos : 1884/1886/1887/1888/1889/		/	/	/
GRados: 03/18/18/30/33/		/	/	/
CarLog: TS /---/GT /OR1/VM /		/	/	/
CauBaj: TRAS		FchBaj: 1889/08/23		
LogAnt: HISPALIS 8				
CarPol: DIPUTADO PROVINCIAL		ParPol: PSOE		
RElig.: CT				
ARchiv: ARCH. HIST. SALAMANCA				

Figura 5

Una vez introducidos los apellidos, el nombre y el simbólico del masón, esto es, los datos que vinculan entre sí dos de los archivos de la base de datos relacional, el de afiliación a la logia y el biográfico, podría darse una de dos circunstancias: o bien es la primera vez que el individuo en cuestión se introduce en la base de datos, o bien un masón con tales nombres, apellidos y simbólico ya se ha introducido con anterioridad en el archivo como miembro de otra logia. En el primer caso el operador podría expresar los datos biográficos del masón, sobre la misma pantalla de introducción de datos; en el segundo, tales datos biográficos aparecerían automáticamente sobre los campos correspondientes y el operador no tendría más que, en su caso, completarlos o corregirlos.

De la misma manera, en cuanto se expresara el nombre de la logia aparecerían en la pantalla los datos relativos a ésta, si ya existían en el archivo por haberse introducido los datos de afiliación de otro masón a la misma logia, y el operador podría entonces modificar o completar los datos; o bien, en caso contrario, esto es, si se trata de una logia de la que nunca antes se ha introducido uno de sus afiliados, el operador expresaría entonces los datos relativos a tal entidad masónica.

Resumiendo, los datos biográficos del masón (esto es, los que figurarían en el que venimos llamando archivo biográfico, figura 4) sólo se expresarían una vez: la primera vez que tal individuo entra a formar parte de una logia o la primera vez (siempre según el orden en que se introducen estos datos en el archivo) que se introduce un miembro de tal logia, respectivamente; es decir, cuando previamente no existían esos datos en el archivo. En ocasiones posteriores no habrá que hacer sino modificar o completar, si fuera el caso, los datos introducidos con anterioridad.

Por las razones aducidas cuando tratábamos de justificar la conveniencia de modificar la estructura del banco de datos de la masonería española se hace especialmente evidente la necesidad de tal modificación en lo referente a disponer de un archivo separado (pero relacionado) de entidades masónicas. La necesidad (incluso la conveniencia) de contar con un archivo biográfico puede sin embargo ser puesta en cuestión. Un estudio profundo, por parte de los especialistas en el tema de la masonería, de las ventajas o inconvenientes de este tercer archivo, así como de los nuevos campos propuestos (e incluso de los que hasta ahora han venido formando parte del registro del banco de datos actual), debe ser un paso previo ineludible antes de afrontar la reconversión definitiva que proponemos.

Finalmente queremos aprovechar la oportunidad para lanzar un par de sugerencias en relación con el tratamiento por ordenador de información referente a la masonería. Estas sugerencias, que dirigimos a la comunidad de investigadores de temas masónicos y que no dudamos que calificarán de interesantes, son, de un lado, la creación de una base de datos bibliográfica que recoja todos los trabajos publicados en todas las épocas sobre la masonería, y de otro, de una base de datos gráfica (aprovechando las facilidades de que disponemos actualmente para la digitalización y manipulación de imágenes) que recoja, describa y clasifique la riquísima iconografía vinculada al tema masónico. Desde nuestra posición de historiadores interesados en el estudio y la explotación de las nuevas posibilidades que nos brinda la informática, esperamos que nuestras propuestas sean acogidas con interés por los estudiosos de la masonería.