

Una aplicación para el manejo de bibliografía e información documental derivada

Jesús Adánez Pavón

Departamento de Historia de América II (Antropología de América). Universidad Complutense. 28040 Madrid.

1. Introducción

Aunque aceptemos que toda aplicación informática —incluso la inteligencia artificial al decir de J.C. Gardin— tiene un carácter instrumental, en la medida que cualquiera de ellas no supone sino un apoyo para la más eficiente realización de tareas posteriores, no cabe duda que unas se aproximan más que otras al proceso final de la interpretación. Si pudiésemos asignar grados de proximidad, dentro de tal escala, a las distintas aplicaciones, es claro que ésta que presento se situaría en el extremo más lejano. El manejo de la bibliografía y de las informaciones extraídas de ella constituye un trabajo previo —de elaboración y clasificación de fichas— que no tiene reflejo directo en las publicaciones y para el que cada cual tiene su propio método. Esta labor, sin embargo, no deja de ser compleja, sobre todo ahora que la producción bibliográfica se multiplica y diversifica y aún más cuando se requiere el uso de trabajos procedentes de distintas disciplinas científicas.

La necesidad de contar con un sistema informático para gestionar el tipo de información documental a que me refiero surgió en este contexto, como parte de una investigación intercultural de la conducta espacial basada en monografías etnográficas publicadas, con el objeto de apoyar y contrastar algunas reflexiones sobre los supuestos y técnicas de la arqueología espacial. Tres eran los objetivos básicos por alcanzar:

1. Mantener relacionados mediante referencias cruzadas tres ficheros: uno bibliográfico, un segundo con las notas literales o comentarios extraídos de la bibliografía y un tercero con datos espaciales normalizados extraídos de los anteriores.
2. Elaborar un tesoro temático y etnográfico válido para la clasificación de los tres ficheros mencionados, de manera que la información pudiera ser seleccionada con criterios diversos.
3. Agilizar las tareas de recuento —elaboración de tablas de frecuencias— y selección de la información.

Estas páginas muestran las características generales de la aplicación que, programada en dBASE IV y bautizada GRIOT —con la esperanza de emular a los conocidos 'archivos parlantes' africanos—, intenta responder a tales necesidades. En primer lugar será descrito el diseño de los ficheros y su estructura de relaciones para, luego, centrar la discusión en el contenido del tesoro y su uso en la elaboración de argumentos de búsqueda, y concluir con un repaso de las distintas operaciones de actualización y recuento

y las posibilidades de manejar los resultados con programas comerciales de tratamiento de textos y cálculo estadístico.

2. Los ficheros y sus relaciones

Existen en el mercado programas creados expresamente para el tratamiento de información documental y, por tanto, bien adaptados a las peculiaridades de este tipo de datos, en especial a la indeterminación de su estructura (*campos*) y de la longitud de los contenidos. Pero todos ellos carecen de la posibilidad de relacionar unidades distintas. Manejar los tres ficheros señalados más arriba con un programa de este tipo requeriría integrarlos en un único archivo donde habría que repetir, para cada ficha de contenido, por ejemplo, su referencia bibliográfica completa.

No renunciar a mantener las relaciones, de acuerdo con el primero de los tres objetivos enumerados anteriormente, y huir de la redundancia en la información, conducen directamente hacia los gestores *relacionales* de bases de datos, como lo es dBASE. Ello exige, por contra, un análisis cuidadoso de la estructura de nuestros datos, en la medida que la mayor flexibilidad en el establecimiento de relaciones está acompañada de una total rigidez en la definición de los campos y sus dimensiones.

De las diversas posibilidades que pueden ensayarse para reducir la redundancia y normalizar la información, la desarrollada por GRIOT se basa en la existencia de cinco tablas distintas:

1. Datos de edición (series, editoriales o revistas).
2. Bibliotecas.
3. Datos bibliográficos, con referencia a las dos tablas anteriores.
4. Notas referidas a la tabla de bibliografía, incluyendo una columna especial de longitud variable (campo *memo*).
5. Otras tablas referidas bien a las notas o bien a la bibliografía, pudiéndose utilizar únicamente una de ellas simultáneamente.

En el caso de la investigación que dio origen a esta aplicación, por ejemplo, se están manejando dos ficheros distintos en el lugar de la quinta tabla. Uno de ellos resume una serie de características de los análisis espaciales arqueológicos publicados y se vincula, pues, con la tabla de bibliografía. El segundo recoge datos espaciales normalizados basados en las notas entresacadas de las monografías etnográficas y plantea una relación más compleja, pues una misma ficha de contenido puede haber sido utilizada en la

elaboración de varias 'fichas espaciales' y, a su vez, éstas suelen resumir datos de varias citas. Para las tablas que han de mantener este tipo de relación (denominada *man*), GRIOT genera una tabla intermedia cuyas filas representan los múltiples emparejamientos de los registros de ambos ficheros (véase CEBRIAN 1988: 106-107).

Para el usuario, todo este entramado se traduce en la posibilidad de manejar siete *vistas* o modelos de ficha: revistas, series, editoriales, bibliotecas, bibliografía, notas y 'otros', a las que se añaden dos más, tesoro y atlas etnográfico, que son comentadas a continuación.

3. El tesoro

Como es bien conocido, los datos que componen una ficha bibliográfica (y, en distinta medida, los de las restantes fichas a que venimos aludiendo) no incluyen información suficiente como para describir de forma aceptablemente completa la realidad que representan (sea ésta una publicación, el análisis de un método o la descripción de una determinada conducta). Por ello es necesario añadirles descriptores que muestren un número de rasgos pertinentes y que, obviamente, han de estar normalizados. La existencia de la Clasificación Decimal Universal (CDU) nos ahorra el esfuerzo que supone la recopilación de una lista de materias y nos ofrece, además, la ventaja de una codificación que permite reducir la longitud de las descripciones. GRIOT se vale de la CDU para organizar su tesoro, pero con dos modificaciones.

La primera de ellas, prevista en la propia CDU, ha sido la introducción de nuevas materias y códigos acordes con la información por clasificar (básicamente etnológica y arqueológica). La segunda es más profunda —hasta el punto que la clasificación resultante deja de ser decimal y universal— y por ello exige una justificación más detenida.

La CDU tiene una estructura jerárquica que parte de 10 grandes áreas de conocimiento representadas por los números del 0 al 9; estas grandes áreas se subdividen en diversas ramas, cada una de las cuales añade un nuevo dígito a la cifra que recibe. Dado que el proceso de subdivisión es autónomo en las distintas ramas, una misma entidad puede tener códigos distintos en función de la disciplina en la cual sea incluida. Por ejemplo, una ficha bibliográfica de una monografía sobre la estructura de determinado asentamiento puede recibir el código 911.37 si se trata de un trabajo geográfico, el 903.3 si es arqueo-

lógico y el 711.4 si está escrito desde la perspectiva de arquitectos o urbanistas; la descripción de un rito funerario puede ser clasificada con el código 393 o con el 903.5 dependiendo de que el rito sea etnográfico o arqueológico.

Puede que este hecho no tenga importancia en una biblioteca cuyos fondos sean consultados manejando fichas de cartulina. Después de todo, este tipo de ficheros ha de organizarse de forma jerárquica, dividiendo sucesivamente las materias. Un fichero informatizado, sin embargo, puede clasificarse de tantas maneras como campos de información incluya sin necesidad de duplicar las fichas mismas. No sé si un ingeniero, consultando acerca de los tipos de barcos balleneros, consideraría útil recibir información sobre las grandes balsas (*umiaks*) desde las que los esquimales cazan la ballena, pero desde luego en el planteamiento que dio lugar a GRIOT —y quizá, aunque en distinta medida, en la organización de cualquier fichero documental en las ciencias sociales y humanas— esta forma de consulta es fundamental.

Por todo ello, GRIOT estructura su tesoro en la forma de una tabla o matriz con 8 columnas o tipos de descriptores: DISCIPLINA, SISTEMA, PARADIGMA, TEORIA, OBJETO, CRONOLOGIA, GEOGRAFIA y ETNIAS. La primera de ellas describe el área de conocimiento a la que se adscribe la información bibliográfica y utiliza la CDU. Las tres siguientes (véase CLARKE, 1972) se refieren respectivamente: (1) al aspecto técnico, económico, social o cultural descrito o analizado ('subsistemas'); (2) al enfoque general del tema ('sociológico' o 'antropológico', 'geográfico' o 'espacial', 'ecológico' y 'morfológico'); y (3) a su ubicación en un determinado tipo de teoría ('deposicional', 'postdeposicional', 'recuperativa', 'analítica', 'interpretativa' y 'meta-teórica'). En la columna de objetos se anotan los elementos materiales envueltos (agujas, templos, asentamientos...), siguiendo la clasificación de Murdock *et al.* (1963), y en las restantes la localización cronológica (períodos o etapas), geográfica (países y áreas culturales) y étnica (según el *Ethnographic Atlas* [MURDOCK, 1967]).

Puesto que las ocho columnas son independientes, las consultas tienen la capacidad de combinarlas de todas las formas posibles. Una consulta sobre, pongamos por caso, punzones de hueso (objeto) desde el punto de vista 'recuperativo' (teoría) en arqueología (disciplina) seleccionaría información referente a cómo excavar este tipo de materiales; una segunda consulta sobre la interpretación (teoría) simbólica (sistema) del espacio (paradigma) obtendría infor-

mación procedente tanto de la arqueología como de la antropología cultural, la semiótica, la arquitectura o la geografía (áreas todas ellas que han considerado adecuado tratar este tema desde su perspectiva).

La clasificación étnica, para terminar, está acompañada de los códigos que describen las distintas etnias en el *Ethnographic Atlas* de Murdock, de forma que las consultas basadas en este descriptor puedan referirse no sólo a un grupo concreto sino también a determinados rasgos culturales (patrón de asentamiento, tipo de agricultura, forma de la casa...).

Hay que convenir que los encabezamientos del tesoro así definido pueden resultar peculiares en exceso. La ruptura con la CDU supone una pérdida evidente de estandarización. Aun más: describir una ficha mediante ocho campos con códigos más o menos extraños requiere un esfuerzo de clasificación que, evidentemente, supone el conocimiento de sus contenidos. No obstante, en GRIOT se ha preferido sacrificar la solución de todos estos problemas en aras de potenciar las formas de recuperar la información.

En cualquier caso, la definición de encabezamientos dentro de la aplicación es una operación abierta: el tesoro es un fichero con materias y códigos al cual se pueden añadir en cualquier momento nuevos registros (borrarlos o modificarlos, sin embargo, requiere una larga revisión de todos los archivos para borrar o sustituir los códigos eliminados). Por otro lado, la clave que conforma el conjunto de descriptores de una ficha puede considerarse, en última instancia, un 'pseudo-número de la CDU' con dos códigos relacionados (disciplina y sistema) y una serie de auxiliares comunes de punto de vista (paradigma y teoría), objeto, cronología, localización geográfica y etnia.

4. Operaciones con los ficheros

Desde la aplicación se pueden realizar las operaciones habituales sobre la base de datos: actualizar la información (añadir, modificar, borrar y recuperar registros), ver las fichas y contarlas. Tales operaciones pueden dirigirse a cualquiera de las tablas descritas más arriba: edición, bibliotecas, bibliografía, notas, tesoro u otros ficheros que eventualmente hayan sido creados en GRIOT. Por supuesto, dada su estructura relacional, la selección de fichas de uno de los archivos puede incluir argumentos referidos al resto.

El engorro que supone construir condiciones de selección cuando los campos son tan numerosos y los descriptores están codificados se ha intentado atenuar mediante un procedimiento que, a través de *menús* desplegables, presenta las distintas posibilidades y permite su libre combinación. A esta rutina se accede desde cualquiera de las operaciones condicionales.

Para el tratamiento estadístico de los datos se han previsto tres formas de cuantificación: (1) recuentos simples de los registros que respondan a un determinado argumento; (2) tablas de frecuencias, resultantes de la intersección de varios argumentos; y (3) matrices nominales, que resultan de la asignación a cada registro de valores alfanuméricos en función de su adecuación a condiciones definidas. Estas matrices pueden manipularse desde programas comerciales de estadística para realizar representaciones gráficas y construir tablas de contingencia sobre las que calcular distintos *tests* y coeficientes. Ello permite, por poner un par de ejemplos, presentar gráficos que muestren tendencias en la historia de la investigación de un determinado tema o medir la asociación entre ciertos rasgos culturales —tal como están definidos en el Atlas— y ciertos tipos de usos del espacio —tal como se han definido en las ‘fichas espaciales normalizadas’—. Como es obvio, la representatividad de la

muestra que sirve de base al gráfico o la fiabilidad e interpretación de asociaciones entre elementos culturales fuera de contexto son cuestiones que ha de evaluar el investigador al margen de la aplicación misma.

BIBLIOGRAFIA

CEBRIAN DE MIGUEL, JUAN ANTONIO

1988 «Hojas de cálculo y programas de gestión de bases de datos», en *Aplicaciones de la informática a la geografía y ciencias sociales* (J. Bosque Sendra *et al.*): 87-110, Editorial Síntesis, Madrid.

CLARKE, DAVID L.

1972 «Models and paradigms in contemporary archaeology», en *Models in archaeology* (David L. Clarke, ed.): 1-60, Methuen, Londres.

MARIN QUIROS, F., A. QUIROS CASADO, y A. TORRES LOZANO

1989 *dBASE IV: técnicas, aplicaciones y rutinas de programación*, Ra-ma, Madrid.

MURDOCK, GEORGE P.

1967 *Ethnographic Atlas*, University of Pittsburgh Press, Pittsburgh.

MURDOCK, GEORGE P., et al.

1963 *Guía para la clasificación de los datos culturales*, Unión Panamericana, Washington.

¿Es posible contar con un centro distribuidor de programas informáticos aplicados a la Arqueología?

José Medina Morales.

Servei d'Arqueologia de l'Institut d'Estudis Ilerdencs.
Apdo. 79. 25080 Lleida.

Actualmente se está dando la circunstancia de la aplicación de la informática por los arqueólogos en la investigación de cualquier tipo de datos que proporciona un yacimiento, muchas veces construyendo sus propios programas y otras aplicando programas que se encuentran en el mercado, consiguiendo algunos unos resultados muy espectaculares, que han hecho que cambie por completo la concepción de los datos que proporciona un yacimiento, pasando éstos de ser considerados inactivos, a ser concebidos como elementos con multitud de variantes, como bien afirma Orton cuando dice:

«El mayor cambio singular que se ha dado en la Arqueología durante los últimos 10 años, especialmente en la Arqueología de Campo, ha sido la difusión del uso de los ordenadores, y sobre todo de los micro-ordenadores. Hace diez años estas máquinas eran principalmente herramientas de investigación, utilizadas en los departamentos universitarios por los teóricos interesados en los enfoques cuantitativos: hoy día forman parte integral de la Arqueología de Campo. Al menos en Gran Bretaña, existen pocos equipos de excavación que no posean o tengan acceso a un ordenador, con una gran variedad de objetivos: proceso de textos, almacenamiento y recuperación de datos, análisis de los mismos, administración financiera, etc... Después de haber cambiado radicalmente la forma en la que los arqueólogos manejan sus textos y datos, el ordenador está empezando a cambiar el enfoque de los registros gráficos: planos y dibujos del yacimiento, dibujo de los objetos, etc. El progreso en este campo ha sido más lento, debido a que la necesidad de tiempo de proceso y de memoria disponible es mucho mayor, pero incluso en esto la utilización del ordenador está pasando de lo experimental a la rutina diaria. A lo ancho del amplio espectro que representa el registro arqueológico, al abaratamiento de los procesadores y la disponibilidad de programas (proceso de texto, Bases de datos y paquetes estadísticos) han cambiado radicalmente la forma de tratar, e incluso de percibir, la información por los arqueólogos». (ORTON, 1988: 13-14).

Como bien dice Clive Orton, estamos en un mundo donde la demanda de información es cada vez más apremiante. No es posible mantener contactos actualizados con toda la literatura especializada, por la cantidad que se genera, no ya en el mundo, sino en la propia península. De esta necesidad nace la idea de creación de un mercado de información para obtener rápidamente una lista de referencias bibliográficas actualizadas, para localizar cualquier información útil, o lo que es lo mismo, crear un banco de datos.