

REVISTA ESTUDIOS

Editorial Alma Mater
1988

Cooperativa Universitaria de Libros
Ciudad Universitaria "Rodrigo Facio"

CONSEJO EDITORIAL

Alicia Vargas de Fournier. Coordinadora

Guillermo Barzuna

Carlos Benavides

Emilia Fonseca

Luis Enrique Gamboa

Enrique Martínez

Hannia Morales

**Coordinación de Edición
Publitex, S.A.**

**Cátedra de Historia de Cultura
Escuela de Estudios Generales
Universidad de Costa Rica**

Lo primero que hay que aclarar al hablar del uso del computador en la investigación histórica es que no constituye un método en sí sino una técnica. De la misma manera, es necesario decir que el computador no sustituye el trabajo del historiador, sólo lo facilita. Lo importante no es tanto la máquina como las preguntas que se le hagan. El uso de computadores interesa desde el punto de vista que permite abordar nuevos problemas utilizando métodos más refinados, particularmente aquellos en los que la cantidad de información requerida es de tal magnitud que su manejo en forma manual se vuelve imposible.

El uso de computador en las Ciencias Sociales se generalizó sobre todo desde fines de los años '50, cuando se pasó del computador de tubos de rayos catódicos a los computadores transistorizados. Estos, conocidos como la segunda generación de computadores, los hizo más pequeños, más baratos y más accesibles.

La tercera generación de computadores apareció en 1963 con el IBM/ 360 (en la UCR hasta el año 81) en el que se introdujeron los circuitos integrados que aumentó la capacidad de almacenamiento y velocidad y disminuyó el costo.

Hasta el momento, la mayoría de los lenguajes de computación (1) tenían una marcada orientación numérica, lo que limitaba el uso

(1) Un lenguaje de computación o lenguaje de programación puede definirse como un conjunto finito de notaciones que sirven para darle instrucciones al computador por medio de un programa. Por ejemplo, para ordenarle al computador que escriba el valor de la variable A, en el lenguaje llamado FORTRAN se escribe:

```
WRITE (6, 100) A
```

```
100 FORMAT (F5.0)
```

y en el lenguaje llamado BASIC:

```
10 PRINT A
```

de computadores para resolver problemas no matemáticos. Esto es comprensible puesto que el computador tenía como finalidad la resolución de problemas especialmente de orden técnico.

En 1965 se desarrolló el lenguaje BASIC, cuyo fácil manejo lo ha convertido en el lenguaje más usado. También en esa fecha empezó a aparecer la preocupación por parte de profesionales de Ciencias Sociales por desarrollar programas (2) que faciliten el manejo y el análisis de información. Este tipo de esfuerzo dio como resultado la aparición de una serie de sistemas integrados de programas de computador orientados al análisis de datos en Ciencias Sociales. De estos paquetes de programas el más conocido y el más utilizado es el SPSS (paquete estadístico para las Ciencias Sociales). El SPSS, así como otros paquetes similares, evita al usuario tener que empezar por escribir sus propios programas pues proporciona una serie de procedimientos fáciles de manejar para la manipulación y el análisis de gran cantidad de información. El uso de estos paquetes y la facilidad de acceso a los computadores que existe hoy, ha generalizado ampliamente el uso del computador en la Ciencias Sociales.

En historia el uso del computador se asocia tradicionalmente a la cuantificación. Si bien los primeros historiadores que lo utilizaron se interesaban sobre todo en problemas cuantitativos (3), hoy su uso va más allá de lo exclusivamente numérico. El almacenamiento y el procesamiento de textos es hoy uno de los más interesantes usos del computador. De la misma manera, se cree que los historiadores que utilizan el computador se interesan únicamente por el análisis estadístico de fenómenos agregados, tales como series de comercio ex-

(2) Un programa es un conjunto de instrucciones escritas en un lenguaje de programación legibles para el computador y en orden secuencial para realizar un trabajo determinado. Por ejemplo, el siguiente programa escrito en lenguaje BASIC calcula la suma de dos variables:

10 READ A, B	lee los valores de las variables A y B
20 C = A + B	asigna a la variable C la suma de A y B
30 PRINT C	muestra el resultado de la suma
40 DATA 50, 60	valores de las variables A y B

(3) Para una revisión de algunas de las primeras aplicaciones del computador a la investigación véase: Edward Shorter, *El historiador y los ordenadores*, Madrid: NARCEA, S.A., 1977, pp. 28-53. Una discusión sobre una de las primeras investigaciones que hizo uso del computador se encuentra en: William O. Aydelotte, "Problems of method in a particular case: the parliament of 1841", en *Quantification in History*, Massachusetts: Addison-Wesley Publishing Company, 1971, pp. 136-154.

terior, crecimiento de la población, resultados electorales. En realidad, la información manipulada por el computador no necesariamente debe ser agregada y anónima.

Una de las aplicaciones no numéricas del computador es precisamente la preparación de listas, inventarios, materiales de archivo, catálogos de bibliotecas o archivos, índices bibliográficos y otro tipo de ordenamiento que permite un acceso rápido y eficiente a la información disponible. Antes de seguir es necesario explicar qué es una base de datos. Una base de datos no es más que el almacenamiento de la información en un medio legible para el computador (discos, cintas magnéticas) y ordenada en una forma que permite un acceso fácil y rápido, por medio de diferentes programas escritos para ello. Por ejemplo, en el caso de una base de datos bibliográfica podemos imaginarla como el almacenamiento del fichero de la biblioteca en una cinta o disco magnético. Cada tarjeta sería un registro que contiene varios campos tales como autor, título, etc. Sin embargo, no sería necesario tener otros registros (tarjetas) por cada título, como ocurre en la biblioteca, ya que mediante los programas de la base de datos podemos relacionar todos los títulos de una misma materia o todos los títulos de un autor sobre una materia específica, o todos los títulos de un autor sobre una materia específica publicados después de 1975, o cualquier otra selección aún más específica.

En el último caso, utilizamos un archivo (fichero) con un tipo de registro (tarjetas), pero una de las ventajas de las bases de datos es poder relacionar registros provenientes de diferentes archivos. Tomemos ahora un ejemplo más complejo, supongamos que queremos relacionar las votaciones en el congreso con las características personales y políticas de los legisladores. En este caso, tendríamos dos archivos (ficheros): uno en el que cada registro contenga el asunto discutido y la votación de cada uno de los legisladores, y otro archivo en el que cada registro corresponda a un legislador y a sus características personales. De esta manera, por medio de la base de datos es posible interrelacionar la información que contienen ambos archivos. Por ejemplo, se puede verificar si todos los legisladores de una misma filiación política votan siempre igual o no, o si existe alguna diferencia en el voto según la edad, o contestar otro tipo de preguntas sin tener que revisar varias veces las tarjetas de los ficheros, que es lo que ocurriría si se hiciera esto a mano.

De hecho, hoy en muchas bibliotecas ya no se buscan los libros por medio del catálogo sino mediante un terminal de computador (4). Igualmente, existen, sobre todo en Estados Unidos, los índices computarizados. De esta manera, es posible tener acceso a una base de datos que contiene información sobre cientos de publicaciones periódicas a lo largo de varios años. Los parámetros de entrada (5) son simplemente de 1 a 5 descriptores, las fechas de las publicaciones y el idioma. Esto permite obtener una bibliografía comentada sobre el tema que se está investigando.

Un segundo ejemplo de aplicaciones no numéricas del computador y una de las más interesantes y complejas es la prosopografía (6). El método más sencillo consiste en realizar un seguimiento de todas las referencias en diferentes fuentes de un individuo particular y sobre la base de estas observaciones reconstruir su ciclo vital o su carrera profesional o política. Hacer a mano este seguimiento en varios documentos resultaría sumamente tedioso; el computador en cambio acelera el proceso y lo hace además más eficiente.

En prosopografía los problemas que se deben resolver en términos de programación son dos: el establecimiento de una base de datos y la interrelación de los registros. Aunque el estudio se refiera al examen detallado de una élite o a un análisis general de una población más amplia, deben definirse una serie de criterios para la inclusión de los individuos y el tipo de información o variables que interesa para cada uno de ellos. Por lo general, se seleccionan una serie de atributos cuantificables aunque también es necesario incluir otro tipo de información, como por ejemplo nombres, lugares, nombres de instituciones y otros. El manejo de esta información numérica y alfabética hace más difícil el trabajo de programación.

Otro problema es que esta información se encuentra por lo general en fuentes diferentes, de ahí la necesidad de establecer

-
- (4) Un computador posee varios terminales, cada uno de los cuales consiste en un teclado y en una pantalla por medio de los cuales el usuario puede comunicarse con el computador. Los terminales del computador pueden estar lejos, incluso a kilómetros de distancia, y se comunican mediante los cables telefónicos.
 - (5) Por parámetros de entrada se entiende los criterios de selección de la información tales como materia, título, autor, etc.
 - (6) De una manera sencilla prosopografía podría definirse como el estudio de algún aspecto de la sociedad por medio de las biografías de diferentes personas que comparten alguna característica determinada.

una base de datos que registre con exactitud la vida de los individuos en toda su complejidad y significación por medio de los atributos y las variables escogidos. La base de datos debe ser suficientemente flexible y sofisticada como para que justifique el tiempo y el costo invertido. Una vez establecida la base de datos el problema que se debe resolver es la interrelación de los registros, tal vez el problema más complejo. Se trata de localizar la información referente a un individuo a lo largo de varios registros diferentes (7).

Un ejemplo de este tipo de aplicaciones es el estudio de Le Roy Ladurie sobre los conscriptos del año 1868 en Francia (8). Utilizando los archivos militares franceses se almacenaron en computador los registros de 11.000 conscriptos que contenían información sobre ocupación, fortuna, alfabetismo, nivel educacional, enfermedades, historia policial durante veinte años, cambios de direcciones de ellos y de sus padres y por último sus características físicas. Esta clase de información resulta así más completa que si se hubiera podido entrevistar a una parte de la población de París en 1868. Algunos de los usos que se le dio a esta base de datos fue el estudio de la delincuencia de acuerdo con categoría ocupacional, nivel de instrucción y otras variables de importancia. Pero quizás la información más valiosa se refiere a movilidad geográfica de la población de acuerdo con distintas variables socioeconómicas.

Otro ejemplo interesante es el desarrollado por Dennis Kelly con los censos realizados durante el siglo XIX en Estados Unidos (9). Este autor almacenó en computador las boletas de los censos de 1850, 1860, 1870 y 1880 correspondientes a las ciudades de Binghamton, New York y Johnstown, Pennsylvania. Esto dio un total de 32.982 registros los cuales contenían 11 variables cada uno.

(7) Véase: J.A. Barman, R.J. Barman y W. T. Kershaw, "Prosopography by computer: the development of a data base", en *Historical Methods Newsletter*, 3 (10), Summer 1977, pp. 102-105.

(8) Emmanuel Le Roy Ladurie, "The conscripts of 1868: a study of the correlation between geographical mobility, delinquency and physical stature, and other aspects of the situation of the young Frenchmen called to do military service in that year", en *The territory of the historian*, Chicago: The University of Chicago Press, 1979, pp. 33-60.

(9) Dennis Kelly, "Linking nineteenth-century manuscript census records: a computer strategy", en *Historical Methods Newsletter*, 2(7), marzo 1974, pp. 72-76.

Las variables se refieren a información tal como nombre, sexo, edad, ocupación, valor de las propiedades poseídas y otras. De esta manera, la base de datos creada permite seguir el cambio en cada una de estas variables a lo largo de un período de 30 años. Resulta evidente que en una investigación de esta magnitud, el computador resulta un instrumento indispensable.

Un uso del computador diferente al de almacenador y clasificador como el que se ha descrito anteriormente, es su utilización como una herramienta sofisticada de análisis matemático. En este campo, la *New Economic History* o econometría retrospectiva, ha jugado un papel tan novedoso como polémico. Lo novedoso estriba en la utilización de modelos lógico-matemáticos y la insistencia en que todas las teorías deben ser expuestas de manera explícita, todos los términos especificados y todas las cantidades definidas. Las ecuaciones del modelo deben incluir todos los términos del análisis. Para el cálculo de estas ecuaciones, el computador se convierte en un requisito indispensable. Lo polémico ha consistido en el prejuicio que ha generado en una gran cantidad de historiadores hacia este tipo de estudios. Esto, por el uso de proposiciones contrafácticas (10) y la falta de crítica de los datos utilizados. Tanto en el conocido estudio de Fogel (1984) sobre los ferrocarriles en la expansión de Estados Unidos (11) como en el de Fogel y Engerman (1974) sobre la esclavitud en el sur de los EE.UU. (12), la rigurosidad del planteamiento formal no se acompaña de una rigurosidad en la selección de los datos. En este caso, no hay nada que el computador pueda resolver: si los datos son pobres, los resultados lo serán igualmente o tal vez más (13).

Otra aplicación del computador y un ejemplo de cómo su uso posibilita el recurso de nuevos métodos es el estudio de las series temporales. Este ocupa un lugar de primera importancia en la histo-

(10) Una proposición contrafáctica consiste en preguntarse qué hubiera pasado si el agente causal de un fenómeno determinado no hubiere estado presente.

(11) Cfr. Robert William Fogel, *Los ferrocarriles y el crecimiento económico de Estados Unidos, ensayos de historia econométrica*, Madrid: Editorial Tecnos, 1972.

(12) R. W. Fogel y S.L. Engerman, *Time on the cross, the economics of american negro slavery*, Boston: Little, Brown & Company, 1974.

(13) David Herlihy, "Numerical and formal analysis in European history" en varios autores. *The New History the 1980s and beyond*. Princeton: Princeton University press, 1982, pp. 115-125.

ria económica. Una serie puede definirse como un conjunto de observaciones en el tiempo, tomadas en forma secuencial (14). El análisis tradicional de una serie buscaba descomponer el movimiento total separando la tendencia, el ciclo, las variaciones estacionales y los componentes aleatorios. Esto se realizaba por medio de técnicas simples como la observación gráfica, el ajuste de una función o el uso de medias móviles, técnicas utilizadas por Labrousse (1944) en su estudio, pionero en este campo, sobre el movimiento de los precios del trigo en la Francia del siglo XV (15).

Con el uso del computador se ha logrado utilizar procedimientos más sofisticados como los modelos explicativos y causales. Así, el método de Box y Jenkins (1970) y el análisis espectral permiten no solo describir la serie sino también encontrar los parámetros que la determinen y aislar las fluctuaciones. Además, por medio del computador es posible el manejo más rápido de series más largas.

En el campo de la historia cuantitativa más convencional, el computador se ha utilizado para clasificar y contar una serie de observaciones relacionadas. El historiador debe definir un conjunto de categorías bajo las cuales van a ser agrupados los datos que contienen los documentos. En otras palabras, debe diseñar un código. Lo que el historiador hace es "encuestar" el documento. Un ejemplo de este tipo de estudios es el análisis de revueltas y disturbios sociales en el siglo XIX hecho por Charles Tilly (1975). Utilizando periódicos, informes policíacos y otro tipo de material descriptivo, contestó preguntas tales como: ¿cuándo ocurrieron las revueltas, cuántas personas participaron, cuál era su origen social, qué daños causaron?, etc. El computador fue entonces utilizado para buscar patrones comunes y correlaciones en los eventos registrados (16).

La modalidad de "encuestar" documentos tiene dos limitaciones. En primer lugar, el cuestionario necesariamente selecciona de antemano la información y, en segundo lugar, no siempre los datos almacenados permiten otro tipo de aplicaciones.

(14) Cfr. C. Chatfield, *The analysis of time series: An Introduction*, London: Chapman and Hall, 1980.

(15) Cfr. Ernest Labrousse, *Fluctuaciones económicas e historia social*, Madrid: Editorial Tecnos, 1962.

(16) David Herlihy, "art. cit", pp. 125-126; también del mismo autor: "Computation in history: Styles and Methods", en *Computer*, 8 (1971), August 1972, pp. 9-17.

Una forma más para utilizar información cuantitativa es aquella en la que el historiador usa las medidas y las categorías ya presentes en los registros (17). Este es el caso en el cual se registran totales de población, precios de las mercancías, precio de los alquileres y otro tipo de información referida a una jurisdicción determinada y a un período finito. Este es el tipo de uso del computador que ha caracterizado al grupo de historiadores franceses que se han identificado en la llamada historia serial. Si bien es cierto que esta clase de estudios empiezan alrededor de los años 30, el computador vino a facilitar el trabajo. Un ejemplo de esto es el estudio de los alquileres en París de 1402 a 1788 (18). La serie se construyó a partir de diferentes tipos de documentos entre los que se revisaron alrededor de 20.000 contratos de arrendamiento. El computador se utilizó entonces para homogenizar y deflacionar la serie.

El tipo de registro que más se presta a la utilización del computador es aquel que ya posee una estructura determinada y que forma parte de una serie o conjunto de registros similares. Ejemplos de esto son los registros de bautizos, matrimonios y defunciones en los registros parroquiales, listas de jefes de familias, registros electorales y otros similares.

El problema es, por lo general, la cantidad monumental de información que hay que manejar: aquí es donde interviene el computador. Para su uso, en este caso, se han seguido tres procedimientos (19):

1. almacenar todo el conjunto de documentos, procedimiento más costoso y complicado;
2. hacer un muestreo del conjunto de documentos;
3. hacer un seguimiento de ciertas variables de interés.

El primero procedimiento es el que tradicionalmente se utiliza en la reconstrucción de familias. Lo que se hace es registrar en forma individual cada acontecimiento (bautizo, entierro, o matrimonio) en una tarjeta. Esto equivale a copiar casi la totalidad de los libros parroquiales. Una vez terminado esto, se reconstruyen las familias: se toma una ficha de matrimonio y se buscan los bautizos

(17) David Herlihy, "art. cit.", 1972, p. 12.

(18) Cf. Emmanuel Le Roy Ladurie, "Changes in Parisian rents from the end of the Middle Ages to the Eighteenth Century", en op. cit., pp. 61-75.

(19) D. Herlihy, "art. cit.", 1982, p. 131.

y los entierros correspondientes y se anota toda la información en una nueva ficha. Si recordamos el ejemplo de la base de datos, no hace falta apuntar las ventajas del computador en un procedimiento como éste.

A pesar de las ventajas que hemos enumerado, existen una serie de problemas comunes que inhiben a muchos investigadores de hacer uso del computador. Un primer problema es que, debido a la falta de formación adecuada, pocos historiadores pueden resolver los problemas asociados con el uso del computador, y mucho menos diseñar sus propios programas. Esto se ve aumentado, además, por los problemas de comunicación con los programadores profesionales. Por otro lado, el acceso a las facilidades de computación de la Universidad no son fáciles puesto que estas están en función de las áreas técnicas. Finalmente, el costo de almacenamiento y procesamiento de datos muchas veces rebasa los presupuestos de la investigación.

Una solución parcial a estos problemas la ha proporcionado el uso generalizado de microcomputadores (20). El microcomputador está constituido por una unidad de procesamiento, el teclado, una o dos unidades de disco y la pantalla (que puede ser un simple televisor). Además, se le puede conectar una impresora. Hoy existen en el mercado una gran cantidad de microcomputadores con características diferentes y que cubren una gama de precios bastante amplia.

La principal ventaja del microcomputador consiste en que no se requiere saber programar para poder usarlo. Existen una serie de programas ya desarrollados, tales como procesadores de palabras, bases de datos y hojas de cálculo, los cuales tienen ya incorporados una serie de instrucciones para el usuario (21).

(20) Véase: Richard Jensen, "The microcomputer revolution for historians", en *Journal Interdisciplinary History*, XVI: 1, Summer, 1983, pp. 91-111.

(21) Un procesador de palabras o editor de texto es un programa que permite al usuario almacenar el texto y luego editarlo. Además de ser mucho más rápido que una máquina de escribir ya no hay que esperar el retorno del carro y los errores se pueden corregir inmediatamente en la pantalla, es posible cambiar de lugar palabras, frases o párrafos, seleccionar el formato de impresión, el tipo de letra, y ejecutar otras funciones de utilidad. Una hoja de cálculo consiste en un programa que almacena información en una matriz, o sea, en filas y columnas. Una vez almacenada la información es posible realizar una serie de operaciones utilizando una o varias celdas de la matriz.

La amplia difusión del uso del computador en la investigación le exige al historiador familiarizarse con este recurso y con las técnicas asociadas a él. Aún a aquellos que opten por los métodos más convencionales, les será necesario poseer algún conocimiento sobre computadores, aunque sea solo para tener acceso a bibliografías almacenadas en computadores o tan sólo para leer los estudios que sí hacen uso de este tipo de técnicas. Para los futuros investigadores, la familiaridad con el computador sí constituirá un requisito indispensable, no solo por la amplia difusión sino porque gran cantidad de las fuentes se encuentran ya almacenadas en computadores.

Finalmente, en ningún momento el papel del historiador se ve disminuido por el uso del computador. Por el contrario, este, a la vez que lo libera de una serie de trabajos rutinarios, también le exige una mayor rigurosidad, sobre todo en lo que se refiere al manejo de los datos. Con datos pobres no se pueden obtener más que resultados pobres. De la misma manera, es al historiador a quien le corresponde escoger y dominar el método, así como formular las preguntas e interpretar los resultados.