

# Metafuente y el uso de los sistemas de información geográfica en historia económica\*

---

**Carrara, Angelo A.; Valencia, Carlos E.; Grava, Massimiliano**

Metafuente y el uso de los sistemas de información geográfica en historia económica\*

América Latina en la historia económica, vol. 25, núm. 3, 2018

Instituto de Investigaciones Dr. José María Luis Mora

Disponible en: <http://www.redalyc.org/articulo.oa?id=279162404002>

DOI: 10.18232/alhe.906

## Artículos

# Metafuente y el uso de los sistemas de información geográfica en historia económica\*

Metasource and the use of geographic information systems in economic history

Angelo A. Carrara <sup>1\*</sup> [carrara@pq.cnpq.br](mailto:carrara@pq.cnpq.br)

*Universidade Federal de Juiz de Fora, Niterói*

ORCID: <http://orcid.org/0000-0002-7807-2632>

Carlos E. Valencia <sup>2</sup> [cvalencia@id.uff.br](mailto:cvalencia@id.uff.br)

*Universidade Federal Fluminense, Brazil*

Massimiliano Grava <sup>3</sup> [cvalencia@id.uff.br](mailto:cvalencia@id.uff.br)

*Università di Pisa, Italy*

América Latina en la historia económica,  
vol. 25, núm. 3, 2018

Instituto de Investigaciones Dr. José  
María Luis Mora

Recepción: 16 Julio 2017  
Aprobación: 26 Septiembre 2017

DOI: 10.18232/alhe.906

CC BY-NC

**Resumen:** Este artículo tiene por objetivo más amplio discutir tanto el concepto de metafuente, acuñado hace tiempo, como el uso de los sistemas de información geográfica en historia. Específicamente buscamos destacar el aumento de investigaciones cuya metodología se asocia al geoprocесamiento o a los sistemas de información geográfica y algunas consecuencias de este procedimiento sobre la historia misma.

**Clasificación JEL:** metafuente, sistemas de información geográfica, historia, historia económica, bancos de datos, B41, C8, N01.

**Abstract:** The purpose of this article is to discuss both the concept of metasource, coined long ago, and the use of Geographic Information Systems in History. Specifically, we wish to highlight the increase in the investigations whose methodology is associated with geoprocessing or geographic information systems and some consequences of this procedure on History itself.

**Keywords:** metasource, geographic information systems, history, economic history, data banks.

The majority of historically oriented websites continue to operate within what has been popularly termed the “Web 1.0” paradigm

Natalie Zacek

## Introducción

En la última década del siglo XX y en las primeras del siglo XXI la investigación en historia ha pasado por un cambio drástico en los métodos que se usan. Este cambio es generado por la transformación radical que han sufrido las fuentes, pues los historiadores se vieron obligados o estimulados, a trabajar con grandes conjuntos de datos.<sup>1</sup>

En muchos casos, se pasó del análisis de algunas decenas, o cientos de documentos, al empleo de miles o centenas de miles de informaciones que provienen de bancos de datos construidos a lo largo de décadas por grandes equipos de trabajo. Esta transformación es bien conocida en el

campo y de su existencia pocos dudarían. Existen trabajos que la describen y explican en detalle (Alves, 2016; Kidner, Higgs y White, 2003).

Uno de los problemas obvios del crecimiento de datos digitales es exactamente su preservación a lo largo del tiempo. De hecho,

once their creators have exploited them to write essays and books, databases tend to be more-or-less forgotten. They are not always updated and sometimes not even preserved; hard disks occasionally crash; and software becomes obsolete. Perhaps most regrettably, the information they contain is not made available to colleagues or to the public. Still, the reality is that many scholars are willing to make their databases available to others once they have used them for their own projects (Dedieu, Marzagalli, Pourchasse, Scheltens, 2011).<sup>2</sup>

Uno de los mejores ejemplos es la base de datos sobre el tráfico atlántico de esclavos que partía de África entre los siglos XVI y XIX (Eltis y Halbert, s. f.) y reunía alrededor de 35 000 viajes y 10 000 000 a esclavos embarcados, con unas cincuenta variables por viaje.<sup>3</sup>

A pesar de que esta transformación es explícita, el uso por los historiadores de esas nuevas fuentes ha llevado a Natalie Zacek (2009, citada por Lichtenstein, 2016, pp. 140-142) a señalar que “la mayoría de páginas de Internet orientadas hacia la historia continúa en el modelo 1.0”. Esto quiere decir que son páginas web que no permiten que la información circule en los dos sentidos, entre servidor y usuario. Son páginas en que la idea de propietario se mantiene en los cánones anteriores a la Internet.

En ese sentido podríamos agregar que no sólo las páginas web mantienen la característica de un generador para múltiples receptores con información que fluye en un único sentido, del primero para los segundos, sino también que buena parte de la producción historiográfica conserva el estilo autor-lector o presentador-auditorio. Como resultado de esa tradición de información que fluye en dirección única, permanecen lejos de los ambientes de los historiadores los artículos escritos por decenas de manos o de trabajo colaborativo entre individuos a miles de kilómetros (algunas excepciones son (Noguerol, Migowski, Giacomolli, Smith, Rodrigues y Pinto, 2007; Spirinelli, 2017). O como señala Jean-Pierre Dedieu (2005, p. 3): “la masse d’information ne#cessaire de#passe les capacite#s d’un chercheur isole# travaillant selon les techniques traditionnelles de la fiche manuelle, mode#le encore dominant dans la profession. [...] le travail collectif est une ne#cessite#”.

El conservadurismo metodológico es fuerte en este medio académico y, en infinidad de casos, el analfabetismo digital no es sólo una norma, sino que parece estar bien visto y ser parte, lo que es paradójico, del conjunto de elementos que proveen el estatus de intelectual para un historiador. De esa forma, el trabajo con millares de datos que provienen de bases disponibles en la web continúa en los formatos de los métodos que usaban relativamente pocas fuentes. Para algunos, es como si no se transformase en nada la investigación, al dejar de estudiar sólo los propios documentos y pasar a utilizar los datos digitales descargados de la web. Sin embargo, debería ser evidente que no es igual, ni aplicable, el método usado cuando se emplean decenas o centenas de datos que el propio historiador consultó

en el archivo o cuando se recurre a centenas de miles, o incluso millones, de datos que provienen de trabajos conjuntos realizados por equipos numerosos.

En ese camino, este artículo tiene por objetivo más amplio discutir tanto el concepto de metafuente, acuñado hace algún tiempo, como el uso de los sistemas de información geográfica (en adelante SIG) en historia. Específicamente buscamos destacar el aumento de investigaciones cuya metodología se asocia al geoprocесamiento o a los SIG y algunas consecuencias de este procedimiento sobre la historia misma.

## El concepto de metafuente

Para comenzar, existe una gran diferencia en el propio concepto de fuente. Las tradicionales, aquellas que son recolectadas una a una por el paciente trabajo del investigador, son bien conocidas. Las nuevas, a las que nos referimos, cuando se usan datos masivos que provienen, con cada vez mayor frecuencia, del trabajo conjunto de grandes equipos, son, como las llamó a mediados de la década de 1970 Jean-Philippe Genet, metafuentes. En aquel entonces, el término se aplicó al ámbito de la prosopografía y su relación con la informática. En términos sencillos significaba una colección de datos científicamente construidos (Genet, 1977, 1986; Maurin, 1982).

Eran los comienzos de la era digital para los historiadores y Genet (1994, p. 9) escribía:

Métasource comme l'ensemble structuré des informations mises en formes et transmises à l'ordinateur et traitées par lui. Tout autant sinon plus que le traitement lui-même, c'est la constitution d'une métasource qui distingue le travail informatisé de celui que ne l'est pas. La métasource, même si elle est constituée d'images ou de sons, ou d'un mélange de ces éléments, qu'elle ait la forme classique d'une base de données ou qu'elle soit constituée de textes au sens le plus classique et habituel du terme, est un texte d'un type particulier, pourvu d'une structure créée par son auteur, et qui, dans tous les as, diffère du texte de la source, même si l'on a affaire à un enregistrement en full-text de la source originelle.

Se observa que un elemento central del concepto propuesto por Genet es la informática, entendida aquí como el procesamiento de datos por computadores. No obstante, si se asume la metafuente sobre todo como fuente construida, entonces el recurso a la informática en las actividades de investigación contemporáneas, a pesar de ser fundamental –y en la práctica imposible de eliminarse– no sería obligatorio.

Como ejemplos de metafuentes pueden ser mencionados, en el caso de la prosopografía, los catálogos o diccionarios de diferentes grupos de personas, pero también las diversas bases de datos (digitales o no) resultantes de la recolección y sistematización de informaciones extraídas de fuentes tradicionales, como por ejemplo, las series de precios o las listas de habitantes.

A esta característica central se le deben agregar dos aspectos importantes enfatizados por Neithard Bulst (1989): *a)* la metafuente debería estar abierta a revisiones y *b)* derivado de lo anterior, es necesario

prestar atención a su metodología de construcción, pues, “once the construction of a database has reached a certain stage it can be very difficult to modify it to meet fresh research demands” (Bulst, 1989, p. 16).

Este aspecto, la posibilidad de revisión de la base de datos es, de hecho, fundamental, porque coloca sobre los hombros de los historiadores la necesidad permanente de atención a la confiabilidad de las bases de datos que se encuentran listas para ser usadas. Se trata, en este ámbito, del viejo deber de siempre verificar si los datos recogidos corresponden, efectivamente, a los que constan en la fuente original.<sup>4</sup> En ese sentido, la metafuente no elimina, de ninguna manera, la necesidad de consultar las fuentes de las cuales se extrajeron los datos que la alimentan. En el caso de las series de precios, por ejemplo, no constituyen metafuentes las series de índices de precios, cuyos valores miden las variaciones del costo de vida de la población con base en el peso de cada mercancía en la estructura de consumo. Como cada grupo social tiene una estructura de consumo particular, el índice producido por cada investigador resulta del conjunto de variables que él haya establecido. Esto es que, en principio, los índices de precios constituyen un resultado de investigación más que una serie de datos de insumo que puedan utilizar otros investigadores sin tener en cuenta los propósitos, las metodologías y las perspectivas teóricas que pueden ser distintas.

Una metafuente es por eso, en términos básicos, un conjunto de datos, sistematizado y relacionado en términos sólo informáticos, que se ha extraído de fuentes diversas, que de ninguna manera ha sido sometido a ningún tipo de estudio.<sup>5</sup> Como las fuentes tradicionales, las metafuentes representan conjuntos de datos brutos. Aunque es claro que, por razones de orden práctico o por facilidad, una serie de datos puede presentarse en formato distinto al original con el objetivo de poder manejarla de forma más simple, sin que esto signifique ninguna alteración a su contenido. Un caso típico de esta situación es la transformación de datos de valores en distintos tipos de moneda colonial a una única serie estandarizada, por ejemplo, de *patacones* a pesos o de *dracmas* de oro a reis. Lo importante es siempre ser completamente explícito en el procedimiento de conversión.

Una expresión que en un primer vistazo parece corresponder a metafuente es *big data*. Aunque no haya consenso semántico, en términos generales, los *big data* incluyen una base de datos compleja (dada por el número de fuentes de datos independientes, cada una con potencial para interactuar; *big*, en este sentido, se refiere tanto al volumen como a la complejidad), valiosa (su valor medido por el grado de innovación que puede producir por medio de técnicas de análisis) y abierta al uso de información longitudinal que complementa el análisis.<sup>6</sup>

Un ejemplo en historia es The Trans-Atlantic Slave Trade Database (TASTD), constituida por decenas o centenas de tipos de fuentes primarias sacadas de archivos esparcidos por todo el mundo, de diferentes tipos (judiciales, comerciales, fiscales, etc.), recabadas por decenas de equipos de investigación y que a lo largo de décadas llenaron el banco de datos bajo la coordinación de un equipo responsable por la estructura del banco de datos, con sus campos y algoritmos. Al final, son más de 35 000 registros

de viajes, con centenas de variables y con consulta en línea. Pero en este caso, se trata de un *big data* que no corresponde a una metafuente, pues la consulta se realiza en un sentido y en un marco definido por los creadores del TASTD, que de hecho combina variables y realiza cálculos de nuevos indicadores, pero que no permite al usuario crear sus propios cálculos.

Un ejemplo trivial puede ilustrar bien el caso. En línea, el usuario puede ver cualquier variable agrupada en el tiempo, ya sea el número de esclavos embarcados en África, o desembarcados en alguna región de América, o la relación de sexo (hombres a mujeres) entre los cautivos; pero el recorte temporal para el cálculo en línea es definido por la red y es anual, quinquenal, decenal, etc., siempre en fechas cerradas, por ejemplo, 1750-1755, o, 1750-1760, o 1750-1775 y no con las fechas que el investigador requiera. Así, el cálculo del agrupamiento temporal de la variable no depende del investigador sino de la red que ya lo trae predeterminado. Claro que el usuario puede bajar los datos y rehacer sus cálculos, pero, en términos en línea, la consulta a la base de datos es en una dirección, que es la que está predefinida por la red.

Otro ejemplo, más interesante y que tiene que ver con los SIG, es que actualmente no existe integración entre los datos del TASTD y una base de datos que determine áreas lingüísticas en el mundo atlántico, con sus modificaciones diacrónicas. Estas dos bases de datos están separadas y diferenciadas y cada una correspondería a un *big data*;<sup>7</sup> esto no implica que estén mal o que tengan que ser corregidas: son buenos ejemplos de *big data*, pero no de metafuentes. Repitamos, un banco de datos aunque provenga de diferentes archivos no es, por tal motivo, una metafuente, mas sí un posible *big data*.

Ahora bien, si un investigador o un equipo de investigadores tomara esas dos bases de datos –áreas lingüísticas atlánticas y TASTD– para producir una nueva base, en ese momento estarían creando una metafuente, pues la estructura de campos y los algoritmos de cálculo debieron ser modificados para integrar dos conjuntos de datos que provenían de ámbitos diferentes. Esa metafuente podrá, con certeza, responder nuevas preguntas, diferentes de las que responden los *big data* de los que procede, puesto que la base permite al investigador crear sus propios cálculos.

No se trata de afirmar que un tipo de base de datos es mejor que otro. Los datos pueden constituir un *big data* o pueden estar en metafuentes, ya que estos dos tipos de conjuntos son diferentes.

En el ejemplo anterior, hay sin duda una similitud entre *big data* y metafuente, porque ambas ensamblan varias fuentes de datos. Sin embargo, como estructura de banco de datos, los *big data* son una única fuente, de tal forma que, si se agregan fuentes primarias, ellas se encajan en la estructura del banco de datos. Así, el banco de datos de los *big data* crece en registros, pero no se modifica. En contraste, la metafuente cruza fuentes y, sobre todo, cruza bancos de datos, con lo que su crecimiento es tanto en el número de registros como en su estructura, esto es en los campos y algoritmos. En ese sentido, los *big data*, en cuanto tales, continúan en la relación de la web 1.0, pues los usuarios alimentan e interrogan al

banco, pero no lo modifican, mientras que la metafuente es más un caso de web 2.0, debido a que no existe propiamente un usuario, pues él puede modificar el banco de datos al ensamblar cosas diferentes y este ensamble queda disponible para los otros.<sup>8</sup>

Otra diferencia importante está en que la gran mayoría de los *big data* son bases de datos que nacen del simple ensamble de datos electrónicos preexistentes, como mensajes electrónicos, videos, informaciones climáticas, señales de sistemas de posicionamiento global (GPS), registros de compra y venta de bienes, por ejemplo. Viceversa, las metafuentes son construidas por los historiadores a partir de fuentes tradicionales variadas. O sea, los *big data* son datos nativos digitales (mono)temáticos y homogéneos relevados de forma instrumental (teléfonos inteligentes, GPS, escáner, transcripción, etc.), como una caja de zapatos cerrada, mientras que las metafuentes son (muchas veces) una colección de datos de diferente naturaleza, y, en general, resultado de la transposición de datos, lo que significa que el investigador puede hacerlos crecer, como una caja de zapatos potencialmente abierta. Además, estos *big data* sirven para realizar análisis descriptivos, predictivos y automatizados; casi nunca se hallan disponibles y, sobre todo, no siempre son compartidos por quienes los producen. La revolución de los *big data* está, hoy en día, en la realización de potentes útiles analíticos (algoritmos) para analizar esos grandes flujos informativos.<sup>9</sup>

Debemos enfatizar que en la metafuente se mantiene la información, sin alteración de los datos, de la fuente tradicional. La diferencia es que la metafuente ensambla diversas fuentes. De esta forma, lo que emerge en, y con, los bancos de datos es una nueva fuente, o una metafuente, ya que está constituida por una serie de partes (o piezas completas) de fuentes tradicionales. Esta característica de ensamble las define como un objeto de trabajo ya, en sí mismo, distinto de aquel que los historiadores hallan en los archivos.

Esta definición trae dos consecuencias importantes para los métodos de investigación. Por un lado, niega que todo trabajo digital, es decir mediado por un computador o sistema informático, sea una metafuente. Esto es, ya que una metafuente surge del ensamble de fuentes, entonces una única fuente que ha pasado por un proceso digital no tiene ningún carácter especial y continúa siendo una fuente tradicional.

Veamos dos ejemplos que representan esta negación de la condición de metafuente, uno obvio y trivial y el otro menos claro a primera vista. El trivial es aquel de servidores en Internet que tienen disponibles fotos, o imágenes, tomadas o escaneadas de las páginas de los libros de archivos (AGN, 2017; Center for Spatial and Textual Analysis (s. f.). En ellas el historiador debe realizar la lectura y el trabajo acostumbrados. La única diferencia con lo que realizaba habitualmente, antes de la era digital, es que no necesita visitar el edificio donde funciona el archivo; en lo demás, las tareas son idénticas. Si bien las copias en la Internet de los libros de los archivos son una tarea importante, como fuente son clásicas y no metafuentes.

El otro ejemplo es el trabajo con imágenes digitales de mapas históricos o las grabaciones digitales de antiguas cintas magnetofónicas. Se trata de aquellos mapas o sonidos que son escaneados o grabados, según sea el caso, en medios digitales para ser cargados en programas informáticos. Si estos mapas o sonidos no son cruzados con otros datos, no pasarán a ser metafuentes, continuarán siendo documentos clásicos.

Estos, en otro lugar, y para el caso de las imágenes, se han denominado mapas estúpidos (Grava, 2016b), pues transforman el soporte material en el que se encuentran originalmente (con frecuencia papel o lienzo) para convertirlos en digitales, pero continúan trabajándose en los moldes típicos, sin que se produzca *nueva* información dentro de la misma fuente. Son estúpidos, ya que no aprovechan en nada todas las posibilidades que los medios digitales les podrían ofrecer, más allá del mero soporte.

En contraste, la condición que tiene la *información* de ser nueva, de no ser explícita, de subyacer sin revelarse de forma obvia, es lo que define a la información de la metafuent, pues el ensamblaje de las fuentes debe llevar a la generación de nueva información, la cual no saltaba a los ojos antes de ser ensamblada con otros datos.

Así, no todo objeto digital es una metafuent, como en el caso de las páginas, sonidos o mapas digitalizados que acabamos de exponer. En síntesis, las metafuentes están compuestas por el cruzamiento de informaciones que generan otras nuevas. Ese cruzamiento emplea medios informáticos ya que procesa millones de datos de forma simultánea. De ese cruzamiento, de ese ensamblaje de fuentes, proviene la generación de información que no era observable directamente en la fuente tradicional.

Siendo así, las metafuentes son, por sí mismas, objetos creativos. No se trata, como en las fuentes tradicionales, de que sólo la lectura de ellas sea creativa. La intención es que las metafuentes ya sean un producto creativo desde el mismo instante en que se construyen. Esto es, que la creatividad se deriva de su constitución, de su condición de ensamblaje. Así, dependiendo de cómo fueron ensambladas las fuentes, será posible, o no, hacer explícita una información en particular. Lev Manovich (1988) afirmaba que la solución narrativa en los tiempos actuales eran los bancos de datos; podríamos decir que las metafuentes están, o deberían estar, en el centro del proceso creativo de la investigación en historia. Si ellas no toman ese lugar, los historiadores continuarán en la era analógica y en el analfabetismo digital, incluso cuando leen sus fuentes en pantallas de computador o cargan sus sonidos o mapas en programas informáticos.

## Objetivo y construcción de metafuentes

Si la metafuent es un acto creativo que determina lo observable y, por lo tanto, define los hallazgos en la investigación histórica, entonces sus objetivos y sus reglas de construcción deben ser explícitos debido a que determinarán su alcance.

En ese sentido, cada tipo de metafuent tiene sus propios objetivos y exige sus métodos de construcción particular. Si, en la definición, toda metafuent es el objeto procedente del ensamblaje de fuentes que revelan

nuevas informaciones que no eran observables, ni explícitas, antes de ser ensambladas y, ese procedimiento de encaje entre fuentes, en la práctica, siempre tiene que ser realizado por medios informáticos, también es cierto que en su construcción cada metafuente puede ser distinta de otra. En otras palabras, en la definición todas las metafuentes son iguales pero en su construcción pueden diferir. Por consiguiente, una metafuente es un acto creativo.

Así, lo que interesa es la construcción de metafuentes mediante el uso de los SIG para la historia económica. Los SIG son bases de datos georreferenciadas que se geoproncesan para producir nuevas informaciones que tienen como referente fundamental el espacio (Mogorovich y Mussio, 1988; Longley, 2011). Para los historiadores estos datos e informaciones se refieren a tiempos pasados y, en general, no observables directamente en el paisaje contemporáneo (Gregory y Geddes, 2014; Owens, Sandes, Zajanc, Stephenson y Dixon, 2014; Bodenhamer, Corrigan y Harris, 2010) y sí mediante las fuentes tradicionales.

Esto significa que todo SIG en historia tiene que tener el carácter de metafuente, porque debe ser el resultado del ensamble de diversas fuentes clásicas, incluyendo mapas, cartografía del pasado o referencias espaciales, que fueron cruzadas con otras fuentes de distinta índole. Ese encaje de fuentes y referencia espacial puede realizarse por medio de programas de computador. Sin duda hay investigadores que emplean cartografía histórica georreferenciada, pero no sistemas SIG ni metadatos. Repitamos, que si es el mapa por el mapa no estaríamos hablando de una metafuente ni de un SIG en historia, sería una historia tradicional, pero en otra plataforma.

Por lo tanto, un SIG en historia es, necesariamente, una metafuente. Si volvemos a la definición de Genet (1994, p. 9) encontramos que: “la métasource est structurée, et cette structure doit en permettre la lecture sous des états variés”. En el caso del SIG en historia como metafuente, esta estructura es un banco de datos, el cual debe permitir la lectura diversa de variables que producen diferentes escenarios.

Estos bancos de datos tienen su propia estructura, lo que de nuevo nos coloca en la interpretación de Manovich (1988, p. 239) sobre las posibilidades y los límites de construcción de esas estructuras:

the general principle of new media: the projection of the ontology of a computer onto culture itself [...] computer programming encapsulates the world according to its own logic. The world is reduced to two kinds of software objects which are complementary to each other: data structures and algorithms. Any process or task is reduced to an algorithm, a final sequence of simple operations which a computer can execute to accomplish a given task. And any object in the world [...] is modeled as a data structure, i. e. data organized in a particular way for efficient search and retrieval.

En consecuencia, lo que parecería ser el problema inicial al comenzar a trabajar con una base de datos, esto es, su estructura en cuanto a captar informaciones, relacionar objetos y procesarlos, que puede ser visto como un mero problema de diseño técnico, de hecho es un problema más complejo, incluso cuando pasa inadvertido, pues la base de datos es la

reducción y, por lo tanto, proyección, del mundo a las dos dimensiones que Manovich señala: estructura de datos y algoritmos (Denley, 1994; Gil, 2015).

Para la metafuente que es SIG en historia, esta condición de representación del mundo es explícita, pues, por un lado, el investigador define cuáles fuentes tradicionales y cuáles informaciones específicas encontradas en ellas va a ensamblar para generar la nueva fuente y, por otro lado, cómo y cuáles referencias espaciales usará para vincular cada registro al espacio físico. Lo que corre ante los ojos del investigador en la pantalla del computador, incluso antes de ser visualizado como un mapa, es un conjunto de datos organizados y una serie de algoritmos establecidos, pero, de hecho, es también un nuevo mundo creado como representación de otro al que está directamente vinculado. Esa creación incluye la dimensión espacial, por tratarse de un SIG, y la dimensión temporal por tratarse de historia. De ello se deriva que la cantidad de información que se desliza por la pantalla en un SIG en historia es un mundo espacio-temporal resultado del acto creativo del historiador en interacción directa con las fuentes tradicionales, de las que provienen los datos recogidos. Ese acto creativo es consecuencia de varios elementos; para comenzar, de la capacidad técnica en el diseño del conjunto de datos y de los algoritmos de la base de datos. Después, depende bastante más de la erudición espacial del investigador y menos de su erudición temporal.

Expliquemos. En general las fuentes informan con claridad del momento específico del que hablan, esto es, de la fecha, año o periodo en que acontecieron los hechos que la misma fuente menciona. Por eso no se requiere mayor erudición temporal para construir la metafuente. El asunto está en la erudición espacial, en el conocimiento de la toponimia antigua: del paisaje de la época y de las formas de nombrarlo, pues las fuentes, necesariamente, referenciarán el espacio usando el código de la época, el cual debe ser traducido para las formas contemporáneas de nombrarlo y de allí para la base de datos. Un ejemplo en México es el *Atlas ilustrado de los pueblos de indios, Nueva España, 1800*, de Dorothy Tanck Estrada, que resultó de la combinación de fuentes modernas y antiguas, a partir de la catalogación de los registros del Archivo General de la Nación (AGN) y del Sistema Nacional de Información Geográfica del Instituto Nacional de Estadística y Geografía, donde contrastó y obtuvo un total de 4 468 localidades y su localización precisa. Además, la base de datos se amplió con los censos de población, las relaciones geográficas y los mapas de la colonia. Como resultado de la integración de estas distintas bases de datos tradicionales, se construyeron mapas históricos con las variables propuestas (Tanck, 2005).

No estamos hablando sólo de los sistemas de proyección espacial y de la asignación de referencias precisas a los elementos que aparecen en las fuentes (Gregory, 2002; Owens et al., 2014; Valencia, 2016a), pues antes de llegar a ese asunto, es necesario que el historiador conozca, sepa, cuáles son los lugares mencionados en la fuente y cómo se denominan esos lugares en el presente.

Esa traducción de la forma de designar los elementos espaciales del pasado para el presente requiere un amplio conocimiento, una erudición al respecto. Como Tiago Gil (2014) comentó, esa erudición fue proscrita durante las últimas cinco o seis décadas y se pretendió desterrarla de la vida de los historiadores profesionales, por considerarla un trazo de la historia vieja y académica. Sin embargo, para los SIG en historia la erudición es imprescindible y tiene que ser rescatada. En palabras de Dedieu (2014):

the greatest challenge facing by now our scientific community, namely the introduction of computing as a basic tool for historical studies [...] Renouncing rigor, documentary critic and erudition is out of question. The rules of historical hermeneutics which our forefathers codified are still valid, and play a central part in our view of what historical computing might be, a point which we shall stress all along. Versatility, I insist, is a master concept in any tool for scientific research.

El empleo de esa erudición espacial, al lado de la incorporación de la dimensión temporal, es lo que diferencia al SIG en historia de los SIG en otros campos del conocimiento. Así, en la construcción de la base de datos, la forma en que se colocará el espacio-tiempo le dará, en primer lugar, la categoría de acto creativo a la metafuente del SIG en historia. De forma más simple podemos decir que una metafuente que es un SIG en historia se construye mediante: *a*) la elección de las formas geométricas que representan los datos de las fuentes, *b*) la capacidad de asignar referencias espaciales a esos datos, *c*) la relación de esos datos con el momento en que ocurrieron y, claro, *d*) la forma específica de organizar los datos. En última instancia, cada SIG en historia es una base de datos en la cual para cada registro hay una referencia espacial, entendida como lugar, una referencia temporal, entendida como fecha, año o periodo, y una referencia a la forma geométrica del registro, es decir, como punto, línea o polígono.

Esta metafuente, como base de datos espacio-temporal, será exportable y otros investigadores tendrán la oportunidad de trabajar con ella. De esta condición emerge una segunda caracterización del SIG en historia como un acto creativo, pues debe ser posible interpretarlo y reinterpretarlo diversas veces, tantas cuantos observadores haya, es decir, cuantos historiadores trabajen con él. Repitamos que aún no hablamos del SIG en historia representado como mapa. Nos referimos a este en su estado de base de datos, de su condición de conjunto de datos y algoritmos.

El acto creativo en la construcción de la metafuente también debe incluir pensar y diseñar cómo se compartirá con otros usuarios. Ya que los otros usuarios deberán, por ejemplo, aumentarla, reducirla, procesarla y, en general, reinterpretarla. Sobre todo, porque esos otros investigadores no necesariamente mantendrán un vínculo directo con los autores. En ese sentido, y para volver al epígrafe de este texto, todo SIG en historia tiene que construirse en los moldes de la red 2.0.

Esto trae al menos dos consecuencias evidentes. Por un lado, la erudición espacial a la que nos referimos como fundamental para la construcción del SIG en historia es condición necesaria, pero está lejos de ser condición suficiente para producir una metafuente. Esto es especialmente importante en estas épocas de recuperación del espacio en la investigación en historia (Guldi, 2011; KüMin y Usborne, 2013;

White, 2010; May y Thrift, 2003), pues, nos parece que, hasta el momento, lo más común no ha sido la incorporación de la variable espacial en la explicación en historia sino la erudición espacial por sí misma. Dicho de otro modo, en buena parte de la recuperación espacial, el espacio continúa sin explicar nada, porque no puede hacerlo si el historiador carece de un problema de investigación.

Por ese camino, la recuperación del espacio parece regresar en la misma perspectiva de la época en que fue abandonada, esto es, porque el espacio no se incorporaba como variable dentro del conjunto mayor de elementos de explicación. Así, la erudición por la erudición no cuenta. Lo que cuenta es la erudición transformada en mecanismo de construcción de la metafuente, como parte del ensamble que la construye.

Por otro lado, la condición de la metafuente como acto creativo que necesariamente incluye, como regla de construcción, su capacidad de interpretación y reinterpretación por otros investigadores diferentes a sus autores, genera que ella esté en la línea de frente contra la visión romántica del trabajo de historiadores que describía Vincent Brown (2016, p. 184): “many historians still subscribe to a monastic ideal of scholarship that emphasizes long years of solitary contemplation, or a romantic ideal, wherein a heroic individual descends into a little-known repository and emerges triumphant with a gift of new knowledge for the world’s admiration. If these archetypes ever approximated reality, few new media projects would develop along such lines.”

El trabajo de esos historiadores eruditos, y sólo eruditos, que se imaginan en esos monasterios en solitaria contemplación para hallar un conocimiento que les permitirá acceder a la admiración general, está lejos, muy lejos, de formar parte del mundo que la construcción de metafuentes y los SIG en historia promueven. No sólo por la falta de afinidad de estas dos concepciones de investigación, sino también porque es virtualmente imposible producir conocimiento con base en metafuentes, si este conocimiento no proviene de una construcción compartida.

Sin duda, en el estado actual de las investigaciones el elemento más débil del universo de los SIG es la existencia (o no) y la disponibilidad (o no) de la información de base que debería ser compartida. Esta cuestión, que no es secundaria, ha contribuido en diferentes lugares del mundo a que el Estado produzca, o se planteé producir, legislaciones que pretenden regular la información que circula sobre paisajes y territorios. Lo que hasta hace unos años (la información sobre los territorios) se desconocía abiertamente o simplemente no se tomaba en cuenta más allá de los especialistas o los interesados directos, como comunidades o agentes locales, ha pasado a ser una de las preocupaciones centrales del Estado.

La entrada del legislador en la regulación de la producción y la circulación de información ha generado estímulos, o por lo menos se propone generarlos, para que el conocimiento sea compartido. La forma de compartir esa información lleva formalmente la expresión de compartir *metadatos*.<sup>10</sup> Estos *metadatos* permiten a los usuarios conocer un conjunto de informaciones sobre los recursos que están utilizando,

como, por ejemplo: los tipos de clasificaciones de la información, la ubicación geográfica de los datos, su refinamiento temporal, la conformidad con los parámetros de calidad y validez, la adecuación a las normas sobre interpolación y las limitaciones al uso y acceso, entre varias otras (*Gazzetta Ufficiale dell'Unione Europea*, 4 de diciembre de 2008, L. 326/12-30).

Este proceso de estandarización, en el caso de Europa, está en ejecución, aunque, es obvio que la velocidad con la que trabaja cada Estado es diferente, lo que vale la pena reseñar son los esfuerzos por estandarizar las formas en que se produce la información geográfica para facilitar su circulación. Sin embargo, y a pesar de estos esfuerzos, todo parece indicar que el desafío mayor todavía se encuentra en la distribución y en las formas de compartir esos metadatos, tanto en la producción directa por parte de los organismos del Estado, como en el caso específico de los historiadores. Puesto que la producción proviene de los especialistas, sean funcionarios estatales, entidades privadas o laboratorios de universidades, sean historiadores o no, el resultado ha sido que la circulación de los metadatos, como observó Paolo Mogorovich (2008): “cozza col modello delle competenze, un modello semplice, utilizzato con una certa efficacia per semplificare problemi complessi, ma che purtroppo ha in sé il seme dell'incomunicabilità”.

Así, a pesar de todo, prevalece la tendencia a ser un modelo construido con cámaras herméticas, en el que cada productor procura controlar la *competencia* sobre sus datos, de tal forma que puede intentar mantener, en la práctica, la información cerrada, ya sea por razones técnicas o simplemente por el cobro del acceso a estos datos.<sup>11</sup> La idea, un poco romántica, del círculo virtuoso de la información por el cual, según Mogorovich (2008, p. 202): “ognuno di noi, nel suo lavoro quotidiano, dovrebbe affrontare i problemi facendo tesoro dell'informazione prodotta dagli altri e, cosa ancora più difficile, una volta prodotta nuova informazione, dovrebbe sistematizzarla e documentarla in modo che gli altri la possano utilizzare”.

Pasa necesariamente por el rígido control de la calidad, incluso cuando se trata de un único investigador. Uno de los problemas centrales en ese esfuerzo por garantizar la calidad, es el de las reglas para *tomar* la información de las fuentes históricas, cuando se trata tanto de fuentes cartográficas como de fuentes textuales. Esto significa definir, explícitamente, que va a dejar de ser información de la fuente para pasar a formar parte de la metafuente.

Estas reglas incluyen, sobre todo, la decisión de cómo manejar los eventuales errores de las fuentes cartográficas o textuales, en especial, de aquellos que tratan sobre elementos geográficos o espaciales, que pueden ser identificados por la diferencia entre lo narrado en las fuentes históricas y lo observable en el presente. El asunto por resolver es: ¿esos errores deben ser anulados a favor de la creación de una metafuente geográficamente correcta en el canon presente?, o bien, por el contrario, ¿ese error debe preservarse para mantener las apreciaciones de la fuente?

Una parte de la respuesta a esta interrogante proviene de la forma de la construcción de la base de datos, que, al igual que cualquier otra base de datos entre historiadores, puede ser diseñada como orientada por la fuente u orientada por el método.<sup>12</sup> Esto da un criterio inicial para resolver el problema. No obstante, otra parte de la respuesta viene de la práctica, siempre que sea viable, de aproximarse lo máximo posible a lo afirmado por la fuente, tal como siempre lo hicieron los historiadores.

Así, el reflejo de los historiadores es mantener el criterio canónico: donde la fuente cartográfica es incorrecta, mantenemos el error. Sin embargo, y sosteniendo que es preciso seguir este criterio, se deben tener en cuenta dos elementos que se contraponen para construir la metafuente. Por un lado, es posible acercarse a la fidelidad con la fuente a través de la referenciación de la información (cartográfica o textual) como un formato *raster*, sin anclarla a una geometría o puntos exactos, esto es, permitiendo que la imagen *raster* sea flexible y fluctúe, dentro de ciertos parámetros, alrededor del espacio. Esto tiene la gran ventaja de no obligar a asignar coordenadas de latitud y longitud precisas a informaciones que provienen de fuentes históricas y que no se dejan anclar en ese formato del presente.

Pero, por otro lado, esta fluctuación y flexibilidad de la imagen *raster* no puede mantenerse sin ninguna referencia al espacio; esta, en algún momento y en algunos *lugares*, deberá estar vinculada o anclada a la geografía. Esto porque el espacio tiene que ser una variable de explicación y el *raster* no puede fluctuar sin ninguna referencia, o porque los programas de computación precisan de esas referencias para poder procesar los datos.

Sobre esa tensión entre flexibilidad y rigurosidad, volveremos más adelante cuando discutamos el tema de los parámetros. Lo que interesa aquí es el tema de la construcción de la metafuente. Como hemos dicho varias veces, las metafuentes se construyen, ante todo, como cualquier otro banco de datos y el investigador debe decidir si empleará un banco orientado por la fuente o por el método. A su vez, esta decisión va a redundar en qué tipo de información adquiere en los documentos (Denley, 1994; Gil, 2015; Vitali, 2004). Ello significa que en ese diseño el historiador ya decide, por lo menos parcialmente, cuál información espacial va a tomar y cómo lo hará.

Este acto creativo del diseño de la base de datos y la cuestión de cuál información tomar de la fuente están indisolublemente vinculados a la tipología de la información con la que se opera. Por ejemplo, cuando se trata de documentos cartográficos y seriales, el ideal de recolección del dato en forma integral puede ser llevado a la práctica. Pero, incluso en un caso así, la información terminará siendo transformada con la incorporación de campos específicos y fundamentales para construir, e interrogar, la base de datos.

Además, los SIG en historia, al igual que todas las humanidades digitales, deben operar con los métodos ortodoxos desarrollados por la informática para que los usuarios puedan cargar información de forma fluida y continua, pero, a su vez, de modo no secuencial (Ragazzini, 2004).

Dejando clara la importancia de esta cuestión sobre el diseño de la base de datos, que debe tener en cuenta la alimentación de la información

de forma no secuencial, es preciso hacer hincapié en que la llegada de la informática, y el consiguiente arribo de fuentes primarias y secundarias digitalizadas, no termina impactando tanto en el proceso de investigación en general, como las formas de pensar, representar y presentar lo que, tradicionalmente, podría ser visto como el producto final y punto de llegada del historiador. La circulación de la información, la construcción –reconstrucción– de los bancos de datos, el cargamento no secuencial de estos datos y el trabajo compartido terminan por presentar el proceso de generar conocimiento de una forma menos lineal y más circular, lo que incrementa la dificultad de mantener la noción de *producto final* inalterado.

Siendo así, el asunto central en torno al cual se ha seguido el debate entre informática-no informática (que una vez resuelto debería contribuir a la incorporación plena de los historiadores en la *villa global*) es si de hecho los historiadores están dispuestos a comunicar los resultados de su trabajo en las formas accesibles que hoy exige la red (Ortoleva, 1996, pp. 81-82).

En otras palabras, se debe garantizar que el “lavoro di verifica e potenziale ricostruzione alternativa [que hoy] è riservato solo a pochi membri della consorteria professionale” debe convertirse, debido a su característica de no secuencial, en accesible, utilizable y deconstruible (Trigari, 2004). Esto debe ocurrir no sólo con el producto, sino también con la divulgación y las formas de compartirlo, que deben estar dentro de los parámetros estandarizados por las formas de *trabajo abierto*, ya que estas definen cómo se comunican los resultados provenientes de producción científica y las fuentes usadas.

Como ya hemos comentado, la migración del dato del archivo (sea cartográfico, serial o textual) para la base de datos es un paso fundamental que implica, de parte del historiador, una selección precisa del método de recolección de la información. El almacenamiento de datos, incluso para fuentes homogéneas como pueden ser las fiscales o catastrales, debe tener, y tiene, siempre la presencia de los parámetros de subjetividad émica – es decir desde el punto de vista en relación con el todo– que la historia digital, en su conjunto, debe siempre tener en cuenta.<sup>13</sup>

## Análisis de metafuentes

Después de que la metafuente se ha diseñado y cargado, con todos los cuidados que hemos señalado y teniendo claro que su carácter no secuencial permitirá continuar con ese ejercicio de forma sostenida en el tiempo, el historiador se enfrenta a su análisis. Antes de dar el siguiente paso debemos aclarar que, de forma bastante sintética, por geoprocесamiento se entiende el tratamiento informatizado de los datos georreferenciados (Zaidan, 2017, p. 7). Los SIG corresponden al conjunto de las herramientas computacionales para el geoprocесamiento, sumadas a la estructura formada por personas, empresas o instituciones, las técnicas y métodos que se articulan mediante rutinas y herramientas programadas en *software* distinto, en aras de

generar datos, principalmente georreferenciados, recolectar, almacenar, editar y procesar estos datos. El objetivo final es generar nuevos datos o informaciones que se difundirán por medio de ese gran sistema constituido.

Si el banco de datos que fue construido es un SIG en historia, tendrá la característica de ser polisémico, de ser leído y releído para buscar múltiples interpretaciones de acuerdo con las formas en que será indagado y observado por el historiador. Así se abre una nueva puerta: la del análisis de la metafuente, la de los mecanismos para explorar la base de datos.

En principio este parece un problema obvio, tanto en su planteamiento como en la forma de solucionarse, pues se trata de la capacidad que tiene un historiador, o un equipo de investigación, de interpretar datos que están referenciados espacial (como lugar y forma) y temporalmente. De esto no cabe duda. Si el SIG en historia ha sido bien construido, entonces será la capacidad de lectura del observador lo que permitirá reinterpretarlo múltiples veces. Sin embargo, la metafuente, por su condición de ensamble de fuentes por medios informáticos, puede generar un peligro al que se debe prestar atención: su supuesta capacidad de exactitud en las respuestas ofrecidas al trabajar con grandes conjuntos de datos.

Como se ha señalado desde hace mucho tiempo (Brown, 2016; Chaunu, 1964; Furet, 1974; Jarausch y Hardy, 1991), cuando se trabaja con grandes volúmenes de datos, que se cargan en sistemas informáticos y se procesan por computador, crece la tentación de suponer que los resultados encontrados están *libres de la contaminación* producida por el investigador. Es decir, se puede tener la tendencia a considerar que el trabajo con una metafuente *libra* al historiador del problema de la subjetividad. Como hemos registrado varias veces, es completamente al contrario, pues la subjetividad está siempre presente en el diseño, construcción, alimentación y –ahora agregamos– el análisis de la base de datos. Aunque, dicho esto, también debemos señalar que la situación inversa tampoco se verifica, es decir, tampoco es realista pensar que los historiadores que trabajan con SIG consiguen, simplemente, manipular sus datos para demostrar aquello, y sólo aquello, que se propusieron demostrar.

En síntesis, ni se puede suponer que el trabajo con SIG en historia elimina la subjetividad, ni se puede pensar que el SIG en historia sea una pura manipulación de datos que obtiene resultados completamente sesgados. Quizá se pueden encontrar ejemplos para los dos casos: tanto para los ingenuos que creen que una base de datos elimina su intervención sobre la información, como para los manipuladores que pretenden demostrar una afirmación sin llegar a investigar, en el sentido pleno del concepto de investigación.

Sin embargo, tal vez la mayoría de las investigaciones se encuentran en un término medio con la intención de ser honestas. O, por lo menos, debemos creer que la mayoría aún busca la honestidad intelectual. De este asunto ya se han ocupado muchos textos. Sobre lo que debemos llamar la atención es sobre una arista particular de la metafuente

que probablemente sea más relevante aquí que en otros casos de la investigación en historia. Se trata del sofisma de la exactitud y precisión.

Como la base de datos es la conjugación de conjuntos de información y algoritmos para procesarlos, sus registros están referenciados espacial y temporalmente, y esos registros son cientos de miles o millones y, por lo tanto, son procesados matemáticamente por medios informáticos, en consecuencia es fácil caer en la tentación de imaginar que los resultados son exactos. Es como si la supuesta precisión que proporcionan los procedimientos matemáticos se elevara de manera exponencial al tratarse de unos datos que están, presumiblemente, referenciados de forma exacta, al espacio y al tiempo. En otras palabras, es pensar que, puesto que el SIG en historia es la representación, como base de datos, de un mundo espacio-temporal, entonces dentro de ese mundo es posible tener precisión y exactitud.

Son casi innumerables las investigaciones que caen en este error y representan con una precisión pasmosa los lugares donde ocurrieron los eventos o, con una exactitud que parece creada por dioses, las regiones donde sucedieron ciertos fenómenos. Así, parece que el investigador y su equipo saben con toda seguridad los puntos, las líneas o los polígonos que determinan y definen los lugares en que sucedieron los elementos o fenómenos investigados.

No obstante, en realidad nunca se tiene la certeza plena de los resultados a los que llega en historia el SIG. Primero, por la naturaleza de la información que proviene de las fuentes que se ensamblan, pues estas proveen la información en el mismo sentido que cualquier otra fuente tradicional (ya que ellas, las fuentes por ensamblar, son eso: fuentes tradicionales) y en consecuencia deben pasar por la misma crítica que siempre deben realizar los historiadores, lo que ya suprime su condición de exacta, como ya lo hemos comentado.

En segundo lugar, que es lo que nos interesa ahora, la referencia espacial asignada a un registro depende de la erudición toponímica del historiador, tal y como lo señalamos, lo que significa que se debe ser completamente consciente de la probabilidad de no identificar, con total seguridad para la escala geográfica en que se trabaja, el lugar del que habla la fuente. En otras palabras, todo banco de datos en historia debe haber sido construido con información filtrada por la crítica de fuente; por lo tanto, nunca será plenamente exacto. Pero, adicionalmente, todos los bancos de datos en historia que fueron georreferenciados tienen por característica específica que la localización espacial no es exacta. Por consiguiente, cada registro en la base de datos contiene información construida por la observación del historiador, ya sea en los campos de lo que la fuente informó o bien en los campos sobre la referencia espacial que apuntan la localización del registro.

Por eso, tal y como lo dijimos antes, el SIG en historia es una clara derivación de la subjetividad del historiador. De aquí su característica de acto creativo. Sin embargo, con esto en mente, no se debe pensar que es pura imaginación. Todo lo contrario, tanto la traducción de la fuente en

información de la base de datos como la ubicación espacial de cada registro implican, y se debe subrayar esto, la objetividad de la investigación.

Esta objetividad se deriva de los límites, claros y explícitos que la realidad determina para la subjetividad del historiador. Así, la observación no es aleatoria, tiene un espectro dentro del cual es posible. Este espectro, a diferencia de la observación directa, sí puede determinarse de forma precisa.

La información, sea espacial, temporal o de cualquier otro tipo, que construye el SIG en historia proviene de la observación y de la erudición del historiador. Esta observación no logra ser exacta, aunque se lo proponga, pero tampoco es completamente aleatoria. De aquí surge una de las grandes fortalezas del SIG en historia, pues, para él, se pueden definir los parámetros que determinan el intervalo en que se hace presente la aleatoriedad.

En detalle, esto significa que la información que proviene de una fuente puede ser georreferenciada asignándole un índice de probabilidad de ocurrencia en un lugar o, también, determinando un área en la que con seguridad sucedió, pero sin definir el lugar exacto. Después, cuando todas esas informaciones se procesan en conjunto, es posible realizar ese proceso asignándole explícitamente los parámetros que se consideraron en la etapa anterior. Por lo tanto, los resultados obtenidos informarán sobre las regiones y las relaciones entre variables, según los grados de probabilidad que existen de ocurrencia (Valencia, 2016b).

Así, se puede saber qué probabilidad hay de que un cierto fenómeno sucediera en un área y cuál es la probabilidad de que en otra área no ocurriera. Esto significa que los análisis de los SIG en historia deberían enfatizar su condición de modelo estocástico y no determinístico. Estocástico en el sentido de ser capaz de reconocer que la incertidumbre está presente, pero que esa presencia puede observarse y formar parte del modelo, o sea que no debe ser excluida ni escondida.

La determinación clara de los niveles de incertidumbre en el análisis enriquece la veracidad de la investigación. En el caso del SIG en historia, al tratarse de una metafuente esa incertidumbre observable es posible, pues se trata de referenciar y procesar inmensos conjuntos de datos. Cada uno de esos datos representa una probabilidad, que al ser agregada en el momento de ser procesada genera un modelo en el cual se establece cuáles relaciones, espaciales o no, tienen más probabilidad de ser relevantes para explicar el problema histórico. La entrada de esta incertidumbre en el modelo, que va a garantizar una aproximación más precisa a la realidad, puede llevarse a cabo de forma intuitiva, a partir del conocimiento del historiador o expresarse formalmente en términos de geoprocесamiento para hallar áreas.

En el primer caso, la investigación de Abreu (2010) sobre los ingenios de azúcar en la Bahía de Guanabara en el siglo XVII puede ser un buen ejemplo. El problema al que esa investigación se enfrentó era comprender la evolución de la producción azucarera en los alrededores del puerto de Río de Janeiro al comienzo del periodo colonial. Para resolver esto, el autor debía encontrar si el número de ingenios había aumentado o

disminuido y si era posible que los incrementos fueran compensados por las reducciones de esas unidades. Por ese camino, era necesario identificar cada una de ellas, lo que resultó ser una tarea imposible, pues las fuentes no permitían tener exactitud en todos y cada uno de los ingenios de que trataban, aunque sí era posible identificar algunos de ellos de forma continua Abreu (2010).

La solución intuitiva de Abreu y su equipo fue definir áreas elipsoidales según su propio criterio. Así, después de décadas de investigación, en las que se cruzaron miles de fuentes, crearon una base de datos que ahora llamamos un SIG en historia. Esas elipses definieron espacios excluyentes entre sí, aunque no abarcaban toda el área del Río de Janeiro rural del siglo XVII, de tal forma que no se tenía una continuidad territorial, al contrario las elipses se asemejan a islas dentro de ese espacio. Después, todo el análisis lo realizaron con base en esas elipses y no por cada uno de los ingenios.

Como ejemplo del segundo caso, cuando se formaliza el cálculo de las áreas de análisis podemos tomar la investigación de Alves y Queiroz (2015), que se ocupó de un problema totalmente diferente: la representación de Lisboa que se hacían los escritores desde mediados del siglo XIX. Para tal efecto se ensamblaron (y se continúan ensamblando, lo que es un buen ejemplo de trabajo continuo y abierto) fuentes en la base de datos. Esta base está georreferenciada y por lo tanto es buen ejemplo de un SIG en historia.

Para establecer cuáles son las áreas representadas por los autores literarios, los historiadores calcularon de manera formal y matemática, a partir de los datos, los mínimos polígonos convexos<sup>14</sup> con 95% de confianza, para las áreas que aparecían en la literatura en cuatro períodos de tiempo. Así, hallaron cuáles espacios estaban siempre presentes en el panorama, cuáles entraron en algún momento determinado y cuáles estuvieron ausentes del horizonte de Lisboa por décadas (Alves y Queiroz, 2015, p. 67).

## Representación de la metafuente

Una de las tantas cosas que deben ser explícitas en la metafuente en general, y en los SIG en historia en particular, es su carácter interpretativo, a saber, en la elección de las fuentes que se ensamblan, en el diseño de la base de datos, en la traducción de la información (espacial o no) que se carga en esa base, en la definición de los algoritmos que relacionan esos datos y en los cálculos que establecen los resultados.

Todo ese conjunto de acciones interpretativas siempre serán, o deberían ser, complementarios para los nuevos investigadores que construyen, destruyen o reconstruyen el SIG en historia, lo que nos lleva a la última característica que debe tener esta metafuente: su interfaz con los reinterpretadores y, de allí, a la representación que hace de los datos, algoritmos y resultados.

Como comentábamos al comienzo del texto, en la investigación con una metafuente el ideal monástico de historiadores que contemplan el

conocimiento en su soledad no podrá realizarse, porque, como hemos insistido, la interacción entre investigadores es un elemento fundamental. Por lo tanto, la interfaz para cooperar y compartir debe formar parte de las preocupaciones de la metafuente.

No es necesario explicar cómo diseñar y construir una base de datos para investigaciones en historia, otros ya lo han hecho (Denley, 1994; Gil, 2015; Vitali, 2004). Es suficiente con enfatizar que las bases de datos tienen una estructura semejante (en un campo, un registro, el tipo de variables, los tipos de algoritmos, entre varios otros elementos) y que siempre deben respetarse para que sea posible que otros investigadores las entiendan y las trabajen (Trigari, 2004). También es importante recordar que las bases de datos deben estar siempre acompañadas de sus cuadernos de códigos, de las explicaciones que las detallan y del diario de trabajo que registra las decisiones que se tomaron en su construcción.

Cuando la base de datos es para la investigación en historia significa que las fuentes fueron traducidas para ser cargadas dentro del conjunto de informaciones que forman la base. Los criterios de esa traducción, esto es cómo fueron leídas e interpretadas las fuentes para ser reducidas a la base, tienen que estar explícitos.

Ahora bien, cuando se agrega a la base de datos la georreferenciación, y ya que se debe asumir que esta es estocástica (aleatoriedad dentro de un espectro definido de posibilidades), se debe tener sumo cuidado para que los investigadores que trabajen con la base sepan de los parámetros de dicha aleatoriedad.

En especial, debe dejarse explícitamente establecida la relación entre escala y rango de probabilidad para la localización de un registro. Es decir, para ciertas escalas existen ciertos márgenes de posibilidad para la ubicación de los elementos. Si un elemento va a ser colocado en un SIG en historia que tiene una escala grande, el margen de localización, en términos absolutos, se reduce, pues en esa escala las diferencias de algunos metros pueden llegar a ser relevantes. Por el contrario, si la escala del SIG en historia es pequeña, el margen de posibilidades para la ubicación del elemento puede ser mayor (Valencia, 2016b).

Debe ser claro que la definición de la escala no es arbitraria y mucho menos definida por la capacidad de construcción del SIG en historia. La escala en la investigación en historia es determinada por el problema que se esté estudiando, es decir, la escala o amplitud geográfica de un fenómeno es una característica del objeto y no una especificidad del método.

En consecuencia, los parámetros que definen los modelos estocásticos para la representación espacial de los objetos deben adaptarse a la escala geográfica del problema investigado y no al contrario. De esto surge un importante desafío: si un SIG en historia tiene que ser interpretado y reinterpretado innumerables veces, según la cantidad de investigadores que lo usen, eso implica que los problemas que debe contribuir a resolver son múltiples, cada uno con su propia escala, lo que deriva en que el SIG en historia debe tener la posibilidad de ampliar o reducir el espectro en el que es posible localizar sus registros.

Dicho de otra forma, ya que el SIG en historia es una representación de la realidad, entonces esa realidad debe contener, en sí misma, la posibilidad de ensancharse o contraerse especialmente dependiendo de la observación, ya que si un investigador necesita, o tiene datos, para representar su mundo en una cierta escala, el SIG en historia, y en especial su interfaz, debe facilitar esa interacción.

De aquí la importancia de la característica de metafuente, pues cada nueva fuente ensamblada podría tener su propia escala de representación que, a su vez, debe *caber* en el SIG en historia. Por lo tanto, la interfaz de alimentación de los datos debe contener diversos campos de localización de los registros dependiendo de la escala que les sea pertinente. Así, no se trata de que cada registro tenga, obligatoriamente, una única localización, pues ese registro puede tener varias, que están vinculadas al espectro de probabilidades en que puede ser ubicado, según la escala en que fue observado. Lo que no deben olvidar los investigadores es que cada localización es pertinente para un rango de probabilidad y una escala.

Es semejante a lo que ocurre con la variable temporal, pues en algunas ocasiones se trabaja con grandes períodos, como décadas o hasta siglos, y en otras oportunidades con semanas o días. Esta escala temporal, tal y como la geográfica, está definida por el objeto investigado. Por eso afirmamos que el mundo representado por el SIG en historia se ensancha o se contrae, espacial y temporalmente, según los intereses del investigador.

Este ensanchamiento y contracción dependen de la interfaz para modificar los campos en la base de datos, de las formas en que esos datos se cargan, pero también del diseño de los algoritmos de procesamiento y de las formas de representación de los resultados.

Esa representación de los resultados debe insistir en el carácter no determinista del SIG en historia. No debe invitar a que los observadores imaginen que el historiador sabe, con precisión, el lugar donde ocurrieron los eventos o con exactitud el peso del espacio en la explicación. Por el contrario, debe poner énfasis en la condición de interpretación que tienen sus representaciones. Lejos deberían quedar los ejercicios de realismo o incluso de ultrarrealismo que aparecen hoy en tantas representaciones de los espacios en el pasado.

Si se insiste en el carácter interpretativo de las representaciones visuales de los SIG en historia se abre un mundo de posibilidades, que van desde los mapas con áreas definidas por manchas de densidad de probabilidad (Grava, 2016a; Valencia, 2016a), pasando por animaciones en las que los objetos animados pueden ser erráticos, vacilantes o que lentamente van abriendo la bruma que oculta regiones (Kahn, Bouie, Martinelli, Salam y Goldberg, 2015), pasando, también, por imágenes que se modifican dependiendo del punto de vista que asume el observador (White, Garton, White y Robertson, 2017) o que alteran objetos en secuencia para insistir en la falta de comprensión que aún tenemos sobre ellos (Brown, 2016, nota 6), hasta llegar a las simples representaciones formalizadas matemáticamente que insisten en los parámetros y relaciones del modelo estocástico.

Cualquiera de las opciones de mapas, animaciones, objetos 3D y expresiones formalizadas para visualizar los resultados de los SIG en historia vuelven a colocarnos en lo que afirmábamos al comienzo: todo SIG en historia es un acto creativo. Aquí, en la representación visual de los resultados, ese acto creativo alcanza su carácter más explícito y pleno.

Al respecto, Brown (2016) expresa:

scholars can advance the pursuit of knowledge by more fully recognizing the qualitative nature of data collection, categorization, and interpretation, by understanding a database as a deliberate, provisional, and even artistic act of historical research. Interface, which mediates between human users and machine protocols, can be designed to highlight rather than obscure acts of interpretation. As it disciplines, constrains, and determines a user's activity, interface then becomes a vehicle for revealing and supporting scholarly interpretations of the assembled archive, and for helping databases to tell stories (p. 8).

Sólo se debe agregar, al lado de esa idea sobre interpretación y provisionalidad, la insistencia en escapar del sofisma del realismo que las representaciones visuales que los SIG en historia pueden traer o sugerir. Así, si lo desconocido, o parcialmente conocido, gana espacio, los observadores serán invitados a participar de las interpretaciones, lo que debe ayudar a romper esa tradicional relación de unilateralidad entre el historiador y su público. Al final de cuentas, de la interfaz, desde el diseño de la base de datos hasta el de su representación visual, depende en buena parte que el trabajo colaborativo en los SIG en historia se fortalezca y se consolide.

## Conclusiones

Así, el trabajo histórico, que comienza sacudiendo el polvo de los archivos, que puede continuar con el estudio sincrónico, diacrónico y comparado de la información producida a través de la metafuente, llega, dentro de ese círculo perpetuo en el que insistimos, al momento de publicar en la red una serie de visualizaciones. Esta es una serie de capas temáticas espaciales que esperan contribuir a generar un conocimiento abierto y, por lo tanto, pensado para ser, cuanto más posible, construido-deconstruido por usuarios que no serán necesariamente expertos.

Nos referimos a la metafuente en su formato de base de datos, que ya hemos comentado. Ahora se trata de insistir en que las visualizaciones de esas variadas capas de información espacial, cada una de ellas representando en el espacio las informaciones, los datos y los vínculos que están en la metafuente, también debe ser abierta y su interfaz debe permitir la interacción de los investigadores y usuarios.

En ese ámbito, los llamados WebMapping tienen un papel crucial, pues ellos permiten que la difusión de este nuevo *conocimiento sea dinámica*, ya que tienen una función clave al poner en marcha redes de contenido en los que es posible, y hasta obligatorio, un papel *activo* para los usuarios en internet en los moldes 2.0.

Las metafuentes publicadas en aplicativos WebGIS o CloudGIS no son simples imágenes utilizadas para ilustrar un fenómeno, tal y como lo

afirmamos al comienzo de este texto. En realidad, estos datos son capas vectoriales y dinámicas a las cuales el usuario puede interrogar en línea consultándolas a través de su navegador y construyendo así su propia visualización o su propio mapa.

La diferencia entre la carta tradicional analógica, a-espacial, en formato de papel y una carta numérica digital, georreferenciada, publicada en WebGIS es, en ciertos aspectos, como la que existe entre una fotografía y una película. La primera información es inmóvil y ofrece una cantidad de información inferior a la que tiene la segunda. En esta segunda, el observador puede percibir una serie de datos y relaciones que no pueden ser hallados en la primera.

El objetivo, cuando llega el momento de publicar las visualizaciones y los resultados, debe ser conservar este elemento dinámico de la información, manteniendo la característica original de espacialización creada en la metafuente que es un SIG en historia. De ahí la importancia de los aplicativos WebGIS que se desarrollaron específicamente para ese fin, es decir que, retomando nuestro ejemplo de comparación, los WebGIS son la plataforma para que la visualización sea como una película y no como una fotografía.

Por lo tanto, además de la *transición digital* de la fuente del estado material, como documento del archivo, a la condición de inmaterial (digital), el nuevo dato en su condición de metafuente deberá publicarse, con todo rigor, en la web (Zorzi, 2000).

El empleo de los aplicativos WebGIS, con los cuales es posible actualizar los resultados y visualizaciones cada vez que se reconstruyen-construyen las metafuentes, permite que los historiadores trabajen en red con grandes bancos de datos referenciados espacial y temporalmente. Esto, como ya hemos dicho varias veces, abre la posibilidad para que los usuarios naveguen activamente por las visualizaciones según sus propios intereses y necesidades.

En resumen, se genera una compleja recodificación del documento que se ha producido gracias a la creación de la metafuente y a la publicación dinámica de los datos, con lo que aparece un nuevo recurso digital que debe circular de forma libre.

## Referencias

- Abreu, M. de A. (2010). *Geografia histórica do Rio de Janeiro (1502-1700)* (Vol. 2). Río de Janeiro: Andrea Jakobsson Estúdio/Prefeitura do Município do Rio de Janeiro.
- Alves, D. (2012). *A República atrás do balcão: Os Lojistas de Lisboa e o fim da Monarquia (1870-1910)*. Lisboa: Cosmos.
- Alves, D. (2016). As Humanidades Digitais como uma comunidade de práticas dentro do formalismo académico: Dos exemplos internacionais ao caso português. *Ler História*, 69, 91–103.
- Alves, D. y Queiroz, A. I. (2015). Exploring Literary Landscapes: From Texts to Spatiotemporal Analysis through Collaborative Work and GIS.

*International Journal of Humanities and Arts Computing*, 9(1), 57–73.  
DOI 10.3366/ijhac.2015.0138

- Bodenhamer, D. J., Corrigan, J. y Harris, T. M. (eds.) (2010). *The spatial humanities: GIS and the future of humanities scholarship*. Bloomington: Indiana University Press
- Brown, V. (2016). Narrative Interface for New Media History: Slave Revolt in Jamaica, 1760–1761. *The American Historical Review*, 121, 176–186. DOI: 10.1093/ahr/121.1.176
- Bulst, N. (1989). Prosopography and the computer: Problems and possibilities. En P. Denley, S. Fogelvik y C. Harvey (eds.), *History and computing II* (pp. 12–19). Manchester: Manchester University Press.
- Carrara, A. (2008). Para uma história dos preços do período colonial: Questões de método. *Locus: revista de história*, 14(1), 187–217.
- Center for Spatial and Textual Analysis (s. f.). *Spatial History Project*. Recuperado de: <http://web.stanford.edu/group/spatialhistory/cgi-bin/site/index.php>
- Chaunu, P. (1964). *Histoire quantitative ou histoire sérielle*. Ginebra: Librairie Droz.
- Dedieu, J.-P. (2005). Les grandes bases de données. Une nouvelle approche de l'histoire sociale. Le système Fichoz. *Revista da Faculdade de Letras. História*, 3(5), 99–112. Recuperado de: <https://halshs.archives-ouvertes.fr/halshs-00004690>
- Dedieu, J.-P. (2014). *Three Pillars of Historical Wisdom: Atomization, Data Building and Flexibility*. Recuperado de: <https://halshs.archives-ouvertes.fr/halshs-00973443>
- Dedieu, J.-P., Marzagalli, S., Pourchasse, P. y Scheltjens, W. (2011). Navigocorpus: A Database for Shipping Information – A Methodological and Technical Introduction. *International Journal of Maritime History*, 23(2), 241–262. DOI: 10.1177/084387141102300211
- Denley, P. (1994). Models, Sources and Users: Historical Database Design in the 1990s. *History and Computing*, 6(1), 33–43. DOI: 10.3366/hac.1994.6.1.33
- Eltis, D. y Halbert M. (s. f.). *Trans-Atlantic Slave Trade*. Recuperado de: <http://www.slavevoyages.org/>
- Frank, A. (2003). Using measures of spatial autocorrelation to describe socio-economic and racial residential patterns in US Urban areas. En D. B. Kidner, G. Higgs y S. White (eds.), *Socio-economic applications of geographic information science* (pp. 146–162). Nueva York: Taylor & Francis
- Furet, F. (1974). O quantitativo em história. En J. Le Goff y P. Nora (eds.), *História: Novos problemas, novas abordagens, novos objetos* (Vol. 1, pp. 49–63). Rio de Janeiro: Francisco Alves.
- Genet, J. Ph. (1977). *Histoire sociale et ordinateur. Informatique et histoire médiévale: Actes du colloque de Rome (20-22 mai 1975)* (Présentés par Lucie Fossier, André Vauchez y Cinzio Violante). Istituto di Storia Medievale de la Università di Pisa/École Française de Rome, Italia.
- Genet, J. Ph. (1986). Histoire, informatique, mesure. *Histoire et Mesure*, 1(1), 7–18.

- Genet, J. Ph. (1994). Source, metasource, texte, histoire. En En F. BOCCHI y P. DENLEY, *Storia & multimedia: Atti del settimo congresso Internazionale Association for History & Computing* (pp. 3–17). Bologna: Grafis.
- Gil, T. (2014). Recuperando terreno: O espaço como problema de pesquisa em história colonial. *Locus-Revista de História*, 20(1), 183–202.
- Gil, T. (2015). *Como se faz um banco de dados (em História)*. Puerto Alegre: Ladeira Livros.
- Grava, M. (2016a). Georeferenziazione e modelli di densità dei mulini a vento maltesi tra XVII e XIX secolo. En C. Valencia y T. Gil (eds.), *O retorno dos mapas: Sistemas de informação geográfica em História* (pp. 257–274). Puerto Alegre: Ladeira Livros.
- Grava, M. (2016b). Imágenes estúpidas versus imágenes inteligentes. Empleo de Webgis y Cloud Services para la publicación de geo-datos. *Revista Uruguaya de Historia Económica*, 6(9), 73–80.
- Gregory, I. (2002). *A place in history: A guide to using GIS in historical research*. Recuperado de: <http://hds.essex.ac.uk/g2gp/gis/sect34.asp>
- Gregory, I. y Geddes, A. (eds.) (2014). *Toward Spatial Humanities: Historical GIS and Spatial History*. Bloomington: Indiana University Press.
- Guldi, J. (2011). *What is the Spatial Turn? · Spatial Humanities*. Recuperado de: <http://spatial.scholarslab.org/spatial-turn/>
- Hayhoe J. (2016). *Geographic and economic factors in internal migration to villages in eighteenth century France*. Chicago: Social Science History Association.
- Inoue, H., Álvarez, A., Anderson, E. N., Owen, A., Álvarez, R. y Lawrence, K.. (2015). Urban Scale Shifts since the Bronze Age: Upsweeps, Collapses, and Semiperipheral Development. *Social Science History*, 39(02), 175–200. DOI: 10.1017/ssh.2015.50
- Jarausch, K. H. y Hardy, K. A. (1991). *Quantitative methods for historians: A guide to research, data, and statistics*. Chapel Hill: University of North Carolina Press.
- Kahn, A., Bouie, J., Martinelli, M., Salam, R. y Goldberg M. (2015). *Animated interactive of the history of the Atlantic slave trade*. Recuperado de: [http://www.slate.com/articles/life/the\\_history\\_of\\_american\\_slavery/2015/06/animated\\_interactive\\_of\\_the\\_history\\_of\\_the\\_atlantic\\_slave\\_trade.html](http://www.slate.com/articles/life/the_history_of_american_slavery/2015/06/animated_interactive_of_the_history_of_the_atlantic_slave_trade.html)
- Kenneth, P. (1967). *Language in Relation to a Unified Theory of the Structure of Human Behavior*. La Haya: Mouton.
- Kidner, D., Higgs, G. y White, S. (eds.) (2003). *Using measures of spatial autocorrelation to describe socio-economic and racial residential patterns in US Urban areas*. Nueva York: Taylor & Francis.
- KüMin, B. y Usborne, C. (2013). At home and in the workplace: A historical introduction to the “spatial turn”. *History and Theory*, 52(3), 305–318. DOI: 10.1111/hith.10671
- Laguardia, R. (2015). *Dos dízimos à demarcação de terras: geoprocessamento aplicado a módulos rurais (Juiz de Fora, séculos XVIII-XIX)* (Tesis de doctorado). Universidade Federal de Juiz de Fora, Brasil.
- Lichtenstein, A. (2016). Introduction. *The American Historical Review*, 121(1), 140–142. DOI: 10.1093/ahr/121.1.140

- Longley, P. (ed.) (2011). *Geographic information systems & science*. Nueva Jersey: Wiley.
- Malavassi, A. P. (2012). Las bases de datos como herramienta para la investigación histórica. *Diálogos Revista Electrónica*, 13(1), 194–198. DOI: 10.15517/dre.v13i1.6370
- Manovich, L. (1988). Database as a genre of new media. *AI & Society. The Journal of Human-Centred and Machine Intelligence*. Recuperado de: [http://vv.arts.ucla.edu/AI\\_Society/manovich.html](http://vv.arts.ucla.edu/AI_Society/manovich.html)
- Maurin, J. (1982). La prosopographie romaine : Pertes et profits. *Annales. Economies, sociétés, civilisations*, 37(5-6), 824–836. DOI: 10.3406/ahess.1982.282906
- May, J. y Thrift, N. (eds.) (2003). *TimeSpace: Geographies of temporality*. Londres-Nueva York: Routledge.
- McFarlane, A. (1978). Computer input of historical records for multi-source record linkage. En *Proceedings of the Seventh International Economic History Conference*, International Economic History, Edimburgo.
- Mogorovich, P. (2008). La provocazione dei Sistemi Informativi Territoriali. *Tra il Dire e il Fare. Notiziario dell'Archivio O. Piacentini*, 10(11-12).
- Mogorovich, P. y Mussio, P. (1988). *Automazione del sistema informativo territoriale. Elaborazione automatica dei dati geografici*. Milán: Masson.
- Mohr, C. O. (1947). Minimum Convex Polygon (MCP). Recuperado de: <http://www.stat.ufl.edu/STA6934/Minimum%20Convex%20Polygon%201%202007.pdf>
- Mukherjee, R. (ed.) (2011). Networks in the first global age, 1400-1800. Nueva Delhi: Indian Council of Historical Research.
- Noguerol, L., Migóski, V., Giacomolli, E., Smith, M., Rodrigues, D. y Pinto, M. (2007). *Elementos da escravidão no Rio Grande do Sul: a lida com o gado e o “seguro” contra a fuga na fronteira com o Uruguai*. Recuperado de: [http://www8.ufrgs.br/ppge/pcientifica/2007\\_02.pdf](http://www8.ufrgs.br/ppge/pcientifica/2007_02.pdf)
- Ortoleva, P. (1996). Presi nella rete? Circolazione del sapere storico e tecnologie informatiche. En S. Soldani y L. Tomassini (eds.), *Storia e computer: Alla ricerca del passato con l'informatica* (pp. 64–82). Milán: Mondadori.
- Owens J., Sandes A., Zajanc C., Stephenson B. y Dixon D. (2014). A geographic information system (GIS): Training manual for historians and historical social scientists. Recuperado de: <https://www.academia.edu/8438126>
- Ragazzini, D. (2004). *La storiografia digitale*. Turín: UTET Università.
- Reips, U. D. y Matzat, U. (2014). Mining “Big Data” using big data services. *International Journal of Internet Science*, 9(1), 1–8. Recuperado de: [http://www.ijis.net/ijis9\\_1/ijis9\\_1\\_editorial.pdf](http://www.ijis.net/ijis9_1/ijis9_1_editorial.pdf)
- Snijders, Ch., Matzat, U. y Reips, U. D. (2012). “Big Data”: Big gaps of knowledge in the field of Internet science. *International Journal of Internet Science*, 7(1), 1–5. Recuperado de: [http://www.ijis.net/ijis7\\_1/ijis7\\_1\\_editorial.pdf](http://www.ijis.net/ijis7_1/ijis7_1_editorial.pdf)
- Sobek, M., Cleveland, L., Flood, S., Kelly, P., King, M. L., Ruggles, S. y Schroeder, M. (2011). Big Data: Large-Scale Historical Infrastructure from the Minnesota Population Center. *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, 44(2), 61–68. DOI: 10.1080/01615440.2011.564572

- Spirinelli, F. (2017). A digital turn? On the future of archives. *DH-Lab*. Recuperado de: <http://www.dhlab.lu>
- Staley, D. J. (2003). *Computers, visualization, and history: How new technology will transform our understanding of the past*. Nueva York: M. E. Sharpe.
- Tanck, D. (2005). *Atlas ilustrado de los pueblos de indios: Nueva España, 1800*. México: El Colegio de México/El Colegio Mexiquense/Comisión Nacional para el Desarrollo de los Pueblos Indígenas/Fomento Cultural Banamex.
- Trigari, M. (2004). *Le riflessioni di Marisa Trigari su la storiografia digitale*. Recuperado de: <http://www.bdp.it/db/docsrv//swf/trigari/definitivo.html>
- Tuckel, P., Schlichting, K. y Maisel, R. (2007). Social, Economic, and Residential Diversity Within Hartford's African American Community at the Beginning of the Great Migration. *Journal of Black Studies*, 37(5), 710–736. DOI: 10.1177/0021934705282376
- Valencia, C. (2016a). *Ao longo daquelas ruas: A economia dos negros livres em Richmond e Rio de Janeiro, 1840-1860*. Jundiaí: Paco Editorial.
- Valencia, C. (2016b). Precisión y exactitud en los sistemas de información geográfica (SIG) en las investigaciones históricas. En C. Valencia y T. Gil (eds.), *O retorno dos mapas: Sistemas de informação geográfica em História* (pp. 223–256). Puerto Alegre: Ladeira Livros.
- Vitali, S. (2004). *Passato digitale: Le fonti dello storico nell'era del computer*. Milán: Mondadori.
- Voutssás, J. (2013). *Cómo preservar mi patrimonio digital personal*. México: Universidad Nacional Autónoma de México.
- White, R. (2010). What is spatial history? Recuperado de: <https://web.stanford.edu/group/spatialhistory/cgi-bin/site/pub.php?id=29>
- White, S., Garton S., White G. y Robertson S. (2017). *Digital Harlem. Everyday life, 1915-1930*. Recuperado de <http://digitalharlem.org/>
- Zaidan, R. T. (2017). Geoprocessamento conceitos e definições. *Revista de Geografia*, 7(2), 195–201. Recuperado de: <https://geografia.ufjf.emnuvens.com.br/geografia/article/view/206>
- Zorzi, A.. (2000). Documenti, archivi digitali, metafonti. *Archivi e Computer*, (10), 274–291.

## Archivos

Archivo General de la Nación (AGN), México. Recuperado de: <http://190.60.123.101:8181/nyssinimag/kwdp/portal/apps/php/protadacatalogoserie.kwe?idioma=ES>

## Notas

\* Aunque el presente artículo sea resultado de un trabajo común, se deben atribuir los apartados Introducción y El concepto de metafuente a Angelo Carrara; Objetivo y construcción de metafuentes a Carlos Valencia, y Análisis de metafuentes y Representación de la metafuente a Massimiliano Grava".

1 Para una presentación somera en cuanto a las bondades y las posibles limitaciones que implica el uso de bases de datos para la investigación histórica,

véase Malavassi (2012). Según esta autora, entre las ventajas de bases de datos electrónicas se cuentan: *a*) unificación de la información, en ciertos casos procedente de diversos focos documentales, en torno a un eje temático específico, ya que una base responde al planteamiento de una problemática muy específica, donde las diversas entradas permiten medir las variables de la investigación mediante sus respectivos indicadores; *b*) automatización y agilización de la consulta de los datos; *c*) maleabilidad al posibilitar la introducción de cambios en la estructura de los datos y la actualización de los mismos; *d*) depuración de los datos con la consecuente eliminación de las duplicaciones.

- 2 Sobre la preservación digital, véase el plan recomendado en Voutssás (2013).
- 3 Otros ejemplos se encuentran en Alves (2012); Frank (2003); Grava (2016a); Inoue, Álvarez, Anderson, Owen, Álvarez y Lawrence (2015); Kidner, Higgs y White (2003); Laguardia (2015); Mukherjee (2011); Staley (2003); Tuckel, Schlichting y Maisel (2007).
- 4 Como ejemplo, la revisión que Georg Welling hizo en la serie original de precios para Holanda organizada por Posthumus, a partir de la cual concluyó que los cálculos no eran muy precisos ni confiables (Carrara, 2008, pp. 173-175).
- 5 Cuando nos referimos a datos relacionados en términos informáticos y no sometidos a un estudio significa que un campo de la base de datos puede estar relacionado con otro en el diseño de la misma base. Por ejemplo, un campo que informa sobre un agente puede estar relacionado con el campo de otro agente a través de relaciones como: “casado con” o “testigo de” o “socio de”, etc. Estas relaciones aparecen en las propias fuentes y no son resultado del estudio del investigador. Sobre el tema de construcción de bases de datos por historiadores véanse, entre otros, Dedieu (2005), Gil (2015) y McFarlane (1978).
- 6 Snijders, Matzat y Reips (2012) muestran que “Big Data is a loosely defined term used to describe data sets so large and complex that they become awkward to work with using standard statistical software”. Por eso “some of the keepers of *big data* sets develop interfaces for everyone to access and analyze some of the data, e. g. Google provides freely available Google Insights, while others hesitate to offer any access”. Más recientemente, Reips y Matzat (2014) mostraron que esta imprecisión terminológica no había cambiado.
- 7 Véase Sobek, Cleveland, Flood, Kelly Hall, King, Ruggles y Schroeder (2011).
- 8 Ejemplo de metafuente, en SIG e historia económica que cruza distintos tipos de fuentes: Hayhoe (2016).
- 9 Un ejemplo de *big data* en línea es <https://www.familysearch.org/search>
- 10 Por esta razón, por ejemplo, el parlamento europeo ha creado (25 de abril de 2007, L108) un organismo supranacional que regula la información geográfica en forma unívoca para todo el territorio comunitario. La infraestructura creada que recibió el nombre de INSPIRE (Infrastructure for Spatial Information in the European Community) nació con la pretensión de “compatibili e utilizzabili nel contesto comunitario e in quello transfrontaliero”. La directiva, 2007/2/CE del 14 de marzo de 2007, el parlamento europeo y el consejo han instituido la infraestructura para la información territorial denominada INSPIRE. El reglamento de actuación realizado por la comisión es el número 1205/2008 del 3 de diciembre de 2008. Esto se encuentra publicado en la gaceta oficial de la Unión Europea del 4 de diciembre de 2008 y en línea en [http://inspire.ec.europa.eu/documents/Metadata/MD\\_IR\\_and\\_ISO\\_20131029.pdf](http://inspire.ec.europa.eu/documents/Metadata/MD_IR_and_ISO_20131029.pdf) [Consulta: septiembre de 2017.]
- 11 Desafortunadamente hay muchos ejemplos de información pública que debe ser pagada, lo que es ya una contradicción en términos. Uno de estos casos es el del Instituto Geográfico Militar en Italia, que a partir de fuentes iconográficas del siglo XIX y comienzos del XX digitalizó y posteriormente vectorizó la toponomástica de esas imágenes. Este proyecto fue financiado con dinero público, pero el acceso a cada topónimo debe ser pagado por cada interesado al *módico* costo de 0.04 euros, lo que, multiplicado por los 700 000 topónimos

- investigados, transforma el proyecto en una forma de recaudo e imposibilita, o impone serias dificultades, a cualquier investigación que necesite grandes conjuntos de datos, como es el caso del trabajo en SIG histórico. Sobre este proyecto de investigación se debe decir que se realizó en una escala 1:25 000, con proyección Gauss-Boaga, encuadrada en el sistema nacional italiano ED40, con coordenadas métricas planas en metros. Véase en <http://www.igmi.org/prodotti/toponimi.php> [Consulta: noviembre de 2016]. Sobre el Istituto Geografico Militare es importante reseñar que en 1861 l’Ufficio del Corpo di Stato Maggiore del Regio Esercito reunió en un único sujeto: l’Ufficio del Regno Sardo, del Reale Officio Topografico Napoletano y dell’Ufficio Topografico Toscano. La nueva institución, que tomó el nombre de Istituto Topografico Militare, fue transferida de Turín a Florencia (1865) y en 1875 pasó a llevar la actual denominación de Istituto Geografico Militare.
- 12 En las palabras de Stefano Vitali (2004), la diferencia entreorientadas por la fuente y orientadas por el método es, de formasimple: “la prima insiste sull’acquisizione integra e integrale della documentazione, mentre la seconda pone l’accento più sui’fatti’, cioè sulle informazioni, più che sulla fonte stessa” (p.15). Aunque existe un espectro de situacionesintermedias.
- 13 Sobre el concepto de émico y su importancia en el diseño de bases de datos, véase Kenneth (1967).
- 14 Sobre los mínimos polígonos convexos (MCP) puede consultarse Mohr (1947).

## Notas de autor

\* E-mail: [carrara@pq.cnpq.br](mailto:carrara@pq.cnpq.br)