



## El diseño de una base de datos para una representación de las fuentes históricas valencianas

### Designing a database for representing Valencian historical sources

CELIO HERNÁNDEZ  
celio.hernandez@uv.es

*Universitat de València*

**Resumen:** En el ámbito de las Humanidades Digitales, es crucial el desarrollo de bases de datos para el acceso y organización eficiente de su contenido, también sobre fuentes históricas. Tanto las «instituciones-memoria» como los proyectos universitarios se especializan desde hace décadas en la creación y mejora de estas, con distintos grados de complejidad, para facilitar la tarea de la comunidad científica y, en ocasiones, centralizar la información que se encuentra dispersa en la red. Esto sucede con el proyecto VaHiFo (Valentiae Historici Fontes), de la UV/IIFV, que, en un esfuerzo por reunir las referencias sobre fuentes históricas valencianas, se ha embarcado en su localización y registro, así como en la creación de un sistema de información que las organice y gestione. El objetivo de este artículo es mostrar el proceso de diseño de dicha herramienta, que atiende a aspectos como la sencillez de uso, perdurabilidad y capacidad evolutiva. Asimismo, se ofrece una metodología que conceptualiza la descripción unificada de las fuentes históricas valencianas, de distinto género y con diversas tipologías materiales o de edición, y que además resalta la relación entre los testimonios de su tradición.

**Palabras clave:** humanidades digitales, base de datos, fuentes históricas valencianas, metodología

**Abstract:** In the field of Digital Humanities, the development of databases for efficient access and organisation of their content, especially regarding historical sources, is crucial. Both «memory institutions» and university projects have specialised for decades in the creation and improvement of these databases, with varying degrees of complexity, to facilitate the work of the scientific community and, occasionally, to centralise information scattered across the web. This is the case with the VaHiFo project (Valentiae Historici Fontes) from UV/IIFV, which, to gather references on Valencian historical sources, has embarked on their identification and registration, as well as the creation of a database to organise and manage them. The aim of this article is to demonstrate the design process of this tool, focusing on aspects such as ease of use, durability, and evolutionary capacity. Additionally, it offers a methodology that conceptualises the unified description of Valencian historical sources, encompassing different genres and various material or edition types, whilst also emphasising the relationship between the testimonies of their tradition.

**Keywords:** Digital Humanities, databases, Valencian historical sources, methodology

DATA PRESENTACIÓ: 5/11/2024 ACCEPTACIÓ: 15/11/2024 · PUBLICACIÓ: 22/12/2024

## 1. Introducción

Hoy en día no es demasiado común que en una revista especializada en estudios culturales y literarios del medievo y la Edad Moderna pueda encontrarse un artículo de amplias características técnicas y que, además, hable sobre bases de datos. No obstante, esto es prueba de cómo los espacios humanísticos paulatinamente se van abriendo a otras perspectivas distintas que acaban redundando en el desarrollo de herramientas digitales que facilitan la investigación y conviven con las metodologías más tradicionales.

Desde la irrupción de las tecnologías digitales en el ámbito académico se ha producido una evolución imparable que afecta al modo en el que se trabaja en los proyectos más tradicionales —las fuentes cada vez son más accesibles, se desarrollan nuevos recursos para analizar documentación literaria, histórica o cultural, etc. Pero también es interesante observar el auge de otras iniciativas que inciden precisamente en la creación de instrumentos y nuevos materiales que favorecen y apoyan otras investigaciones. Este es el caso de aquellas centradas en desarrollar repositorios de información, en diversos grados de complejidad, que acaban revelándose como elementos utilísimos para la comunidad científica.

Respecto a este tipo de herramientas, podemos observar cómo las «instituciones de la memoria» llevan décadas preocupándose por los desafíos y oportunidades de las TIC (Jacinto y Balsa 2008), y lo han hecho compatibilizando sus tareas con el desarrollo digital. Esta situación ha permitido la creación de portales con catálogos o exhibiciones de libros y documentos virtuales para una consulta más accesible y dinámica. No obstante, cada institución, suele dedicarse a poner en valor sus propios fondos y a pesar de que, como veremos, hay propuestas para establecer nuevos vínculos que favorezcan la interpretación de su contenido (Agenjo 2004), las fuentes de los estudios académicos se encuentran diseminadas. Es, por tanto, relevante la labor de grupos y centros de investigación que concentran recursos para aglutinar esta información y reunirla, generalmente en plataformas de datos. Entre algunos ejemplos que se tratan en el presente artículo, debemos destacar el proyecto VaHiFo (*Valentiae Historici Fontes*), del Institut Interuniversitari de Filologia Valenciana y desarrollado en el seno de la Universitat de València, coordinado por Eulàlia Miralles y Marco Antonio Coronel de los departamentos de Filología Catalana y Filología Clásica de dicha universidad. Entre sus objetivos se encuentra el de desarrollar un banco de datos que sirva para recoger la información diseminada sobre las fuentes históricas valencianas, que no atiende solo a los textos originales —manuscritos e impresos— sino que también se ocupa de sus ediciones —históricas, críticas, facsímiles, reproducciones parciales, etc.

Esta iniciativa surge en un momento oportuno donde se puede partir de referentes adecuados y en el cual los proyectos en Humanidades Digitales cada vez son más abundantes, pero con la precaución de saber que en ocasiones estos desarrollos con componentes tecnológicos son de difícil gestión —bien por su coste, o bien porque se requieren conocimientos especializados avanzados para poder ejecutarlo. Como consecuencia de esto, se propone un método de creación

de una base de datos sencilla, pero con una metodología de conceptualización profunda de dichas fuentes para poder ser registradas de forma rigurosa y eficaz. El procedimiento aquí explicado no es habitual entre otros que se pueden encontrar en publicaciones más tradicionales, pero es adecuado para un trabajo interdisciplinar de semejantes características

Este artículo no pretende entrar en los resultados del propio proyecto —todavía en vigencia en el momento de escribir este texto y cuya labor merecería atención en otra publicación— sino ofrecer un estudio de caso de construcción específica dedicada a este tipo de recursos. Asimismo, su relevancia radica también en que existen escasas publicaciones, con perspectiva humanística, sobre elaboración de repositorios de información aplicados en este ámbito. Ejemplos de trabajos que prestan atención a estos sistemas son López (1998), un temprano estudio sobre documentos y bases de datos históricas del momento, la aproximación que hace Malavassi Aguilar (2012), o ya, más en línea con lo que se pretende aquí, la creación y conceptualización de una plataforma sobre divisas históricas de López Poza (2017; 2020).

## 2. Objetivos

Como se anticipa en la introducción, el objetivo principal del presente artículo es mostrar la forma de diseñar e iniciar el desarrollo de un depósito de información sobre todo tipo de fuentes históricas valencianas. Esto se ha hecho con la intención de establecer un paradigma al que se pueda acudir para elaborar iniciativas similares. Para lograrlo todo ello, se han marcado otros propósitos secundarios

Se busca ilustrar la necesidad de por qué se ha llegado a este punto, gracias a una breve visión diacrónica sobre la accesibilidad a otros testimonios históricos mediante bases de datos, hasta llegar a la que aquí nos ocupa, la del proyecto VaHiFo. Este, con la finalidad de crear un repositorio virtual, está elaborando una web —en desarrollo— donde se pueda registrar la información de las fuentes principales, tanto originales como editadas, útiles para los estudios académicos históricos, literarios y culturales escritas en cualquier lengua utilizada en el territorio valenciano desde la Época Clásica hasta el siglo XVIII.

Adicionalmente, se explica la construcción de la base de datos de VaHiFo. Esta ha tenido en cuenta diversos factores en su elaboración, que se analizan a lo largo de la publicación. Basta decir por ahora que este sistema no pretende ser solo un mero contenedor de registros básicos, puesto que en desde su diseño se preocupa por dotar de sentido conceptual a los datos de las fuentes al reflejar las relaciones entre sus testimonios y ediciones académicas. Además, está concebida de forma básica para facilitar su uso y mantenimiento y para que se pueda actualizar a distintos formatos. Se presenta, por tanto, como un recurso imprescindible y valioso para cualquier tipo de investigación que necesite su consulta. Por todo esto, para su desarrollo se han establecido los siguientes objetivos:

- Informar sobre las fuentes.
- Describir los testimonios o ediciones de las mismas.
- Representar la relación entre materialidades y la obra fuente, en caso de ser conocida.
- Ofrecer, si existe, el recurso digital o el enlace al objeto virtual.
- Concebir una sencillez técnica a la hora de completar la base de datos.
- Posibilitar la reutilización de los registros.

Por último, se establece un método que comprenda los propósitos del sistema de datos. Asimismo, ha sido necesario desarrollar una metodología capaz de comprender las relaciones entre los elementos del sistema y transformarlos en registros. Esto, como veremos en el apartado de su diseño, se ha realizado mediante un sistema propio de identificación y clasificación que tiene en cuenta dichos vínculos entre testimonios.

### 3. Las bases de datos y las fuentes históricas

Antes de comenzar a hablar de la conceptualización y diseño de la base de datos sobre fuentes históricas valencianas de VaHiFo, es conveniente establecer una brevísima diacronía sobre la relación entre repositorios de información y fuentes. La necesidad de ello no es solo la de observar su origen sino, más bien, entender la evolución de esta simbiosis poder vislumbrar las actuaciones que desde las Humanidades Digitales se pueden ejecutar con perspectiva innovadora.

A finales de la centuria pasada, Armando Petrucci (1999: 6-7) analizó cómo la cultura escrita ha cambiado a lo largo del tiempo. Esta se ha organizado de dos maneras: «virtual», mediante herramientas como catálogos, compendios, bibliografías y listados; y «real», a través de la conservación física y la disposición sistemática de los documentos escritos. Según Petrucci, esta dualidad organizativa surge de mecanismos de selección y exclusión que reflejan influencias tanto funcionales como ideológicas de las élites intelectuales. Basándose en estos principios, las sociedades han desarrollado estructuras «virtuales» que actúan como nodos esenciales para la consulta, lo que dirige a quien investiga hacia la documentación necesaria para la extracción de conocimiento relevante. Con la llegada de la tecnología, el campo de las Humanidades ha experimentado una rápida transformación en la organización de la información, facilitando el acceso a través de bases de datos y promoviendo una democratización del conocimiento para una sociedad tecnológicamente alfabetizada.

Es de esta manera como los sistemas de información se han convertido en uno de los cimientos de las «instituciones-memoria» en sus objetivos de organización y puesta en valor de sus fondos. Tal y como Xavier Agenjo (2004) afirmó, los profesionales del sector han desarrollado una clasificación sistemática de los registros de objetos mediante el uso del formato XML. Este estándar facilita

la definición de Document Type Definitions (DTD) y Esquemas, herramientas esenciales para la descripción, organización y conexión de datos y que, además, mejora la estructuración de los datos y metadatos, de forma que optimiza su intercambio y transferencia. Así pues, la organización de la información responde a normativas internacionales establecidas por entidades como la International Federation of Library Associations and Institutions (IFLA), el International Council on Archives (ICA) y el International Council of Museums (ICOM). Dichas reglas permiten la aplicación de modelos de catalogación específicos adaptados a las diversas formas del patrimonio cultural.

En la actualidad disponemos de excelentes portales de bibliotecas y archivos cuyas bases de datos desempeñan un papel crucial. En su función de contenedores de información no solamente disponen de los registros de las fuentes históricas necesarias para la labor de la investigación, sino que también nos ofrece la posibilidad de acceder a facsímiles digitales para poder consultarlas de una forma rápida y eficaz en la mayoría de los casos. Ejemplos de sobra conocidos son el portal de la Biblioteca Valenciana Nicolau Primitiu (BVNP) y su Biblioteca Valenciana Digital (BIVALDI)<sup>1</sup> que contiene la reproducción digital de buena parte de sus fondos; en ámbitos nacionales se debe mencionar la Biblioteca Nacional de España (BNE) y su Biblioteca Digital Hispánica (BDH),<sup>2</sup> la Bibliothèque nationale de France (BnF) con su biblioteca digital Gallica<sup>3</sup> o la British Library<sup>4</sup>. En el ámbito académico debemos destacar los portales de las bibliotecas universitarias, como es el caso de las públicas valencianas. Tampoco debemos olvidar el campo de la archivística donde PARES<sup>5</sup> a nivel estatal o SAVEX<sup>6</sup> a nivel autonómico nos garantizan un acceso cada vez mayor y otros como el portal del Arxiu de la Catedral de València<sup>7</sup> donde los esfuerzos de accesibilidad digital a la información de las fuentes son cada vez mayores.

Todo ello genera un cosmos de portales que, si bien nos brinda una mayor oportunidad de consulta de fondos, también genera una sensación de diseminación de la información que refleja la realidad de la situación del patrimonio bibliográfico y documental. No obstante, la evolución tecnológica está teniendo un papel relevante en las soluciones a este problema.

Es en este contexto en el que Agenjo (2004) ya anticipaba el éxito de un elemento que ha modificado sustancialmente el acceso a su contenido en los repositorios de información, la web semántica,

---

1 <<https://bivaldi.gva.es/>> (Consultado el 09/09/2024).

2 <<https://bdh.bne.es/>> (Consultado el 09/09/2024).

3 <<https://gallica.bnf.fr/>> (Consultado el 09/09/2024).

4 <<https://www.bl.uk/>> (Consultado el 09/09/2024).

5 <<https://pares.cultura.gob.es/inicio.html>> (Consultado el 09/09/2024).

6 <<https://savex.gva.es/>> (Consultado el 09/09/2024).

7 <<https://archivocatedraldevalencia.com/>> (Consultado el 09/09/2024).

definida por el Wide Web Consortium como una herramienta que «proporciona un marco común que permite compartir y reutilizar los datos».<sup>8</sup> El autor señaló que las «instituciones de memoria» en su versión virtual, se encontraban en disposición de pasar a la web semántica y, por esta razón, a la consolidación de los datos relacionados haciendo uso de las denominadas «ontologías» o, también, «vocabularios». Todo esto lo defendía como posible debido a la clasificación y organización de contenidos, que se realiza de forma metódica por las instituciones y permite la aplicación de lenguajes y normativas XML y, sobre todo, RDF —Resource Description Framework. Todo esto facilita tanto la descripción como la vinculación de la información y posibilita la interconexión de bases de datos para lograr, incluso, búsquedas conceptuales que además de palabras clave ofrezca información adicional —en el caso de autores notas como nacimiento, muerte, obras, personajes relacionados; en el de libros, ejemplares, ediciones o filiaciones de estas, por ejemplo. A nivel nacional y con la atención puesta en estudio de fuentes históricas debemos destacar el portal DatosBNE<sup>9</sup> en el ámbito bibliotecario y en el archivístico, la asignación de identificadores permanentes a los registros de autoridad del Censo-Guía de Archivos de España e Iberoamérica<sup>10</sup> que busca la interoperabilidad con otras bases de datos. En un nivel europeo debe destacarse indudablemente el papel de Europeana, un macro-portal alimentado a través de distintas instituciones nacionales e internacionales que permiten acceder a un extenso patrimonio digital y digitalizado de todo el territorio transnacional y en varias lenguas, entre las que se incluyen el catalán o el euskera.

Dentro de este panorama, surgen también proyectos académicos basados en sistemas de información cuyas funciones varían entre un espectro más básico, como simples herramientas organizativas para facilitar estudios internos, hasta horizontes más amplios que cuentan con contribuciones de significativo impacto, resultado de años de investigación meticulosa. El abanico es tan amplio que a la hora de diseñar un repositorio para un proyecto se han de tener muy en cuenta las particularidades de los registros a recoger.

Un ejemplo de gran proyecto universitario en el que claramente se aprecia su labor de unificación de la información relacionado con fuentes históricas, aunque en este caso teatrales, lo hallamos en el ámbito de la filología hispánica, el denominado ASODAT,<sup>11</sup> que fue coordinado inicialmente por Teresa Ferrer Valls —entre 2016 y 2023— de la Universitat de València y en la actualidad, por Alejandro García Reidy de la Universidad de Salamanca. En su portal se aglutinan y homogenizan distintas bases de datos pertenecientes a diversos grupos de investigación que estudian las fuentes del Teatro Clásico Español desde el siglo XVI y que analizan diversos aspectos tales como los intérpretes, las noticias sobre representaciones, las censuras y licencias, las canciones o la escritura

---

8 La cita es una traducción propia. Disponible en: <<https://www.w3.org/2001/sw/>> (Consultado el 06/09/2024).

9 <<https://datos.bne.es/>> (Consultado el 09/09/2024).

10 <<http://censoarchivos.mcu.es/>> (Consultado el 09/09/2024).

11 <<https://asodat.uv.es/>> (Consultado el 09/09/2024).

de los manuscritos. El proyecto se presenta como una federación de sistemas de información que estandariza la información de los diversos grupos a través del sistema de *software* libre MySQL de bancos de datos (Ferrer 2018), gracias a ello se pueden interconectar las búsquedas, a través de un solo motor, relacionadas con un autor o una obra.

En el ámbito de la recopilación de fuentes literarias históricas, podemos destacar otro desarrollo distinto de base de datos en el portal NISE —Literatura Catalana de l'Edat Moderna—<sup>12</sup> del Institut de Llengua i Cultura Catalanes de la Universitat de Girona. En este sitio web se ofrecen instrumentos de búsqueda de información y disposición textual de obras de época moderna en lengua catalana, así como investigaciones relacionadas con la temática. Este banco de datos —y biblioteca digital— se establece como una valiosa herramienta académica para favorecer investigaciones fundamentadas y rigurosas sobre la literatura catalana de los siglos XVI y XVII.

Estas iniciativas, que han conseguido objetivos importantes en el campo académico, son solo una muestra de las capacidades del uso de sistemas de datos y proyectos humanístico, además de un referente para nuevos planteamientos. En este aspecto, es posible entender el surgimiento de VaHiFo, que al observar una dispersión significativa en el ámbito de las fuentes históricas valencianas —sin entrar en propuestas editoriales como la de la colección *Fonts històriques valencianes* del Servei de Publicacions de la Universitat de València— y al comprender la necesidad de unificar estos recursos han motivado la creación de una pionera en este ámbito. Compuesto por un nutrido grupo de filólogos de las áreas de Filología Catalana y Filología Clásica, este busca compilar información crucial sobre fuentes para el estudio de la literatura, historia y cultura valencianas y hacerla accesible a la comunidad científica mediante una base de datos. A pesar de que la relevancia de este proyecto es indiscutible, tratarlo en detalle no es el asunto que compete a este estudio sino, más bien, profundizar en la creación de su sistema de información. Para ello, ha sido necesario considerar varios factores técnicos y conceptuales esenciales para desarrollarlo y que cumpla con los objetivos y características específicas de VaHiFo, aspectos que se exploran a continuación.

#### 4. Una conceptualización organizativa para las fuentes históricas valencianas

Como ya se ha mencionado, en el campo del estudio de fuentes históricas valencianas, VaHiFo se destaca en la recopilación de información de las mismas al crear una herramienta de referencia. Entre los objetivos del proyecto se pretende localizar, estudiar y poner al servicio de la comunidad las fuentes halladas a través de su banco de datos, de tal forma que estas puedan ser fácilmente encontradas e incluso ofrecidas virtualmente.

No obstante, el tratamiento de dichas fuentes, por su idiosincrasia, no puede quedarse en una simple recopilación catalográfica. Los estudios ecdóticos realizados a través de décadas de investigación

---

12 <<https://www.nise.cat/>> (Consultado el 09/09/2024).

merecen una especial atención ya que han filiado y vinculado a través de diversas metodologías, como la neolachmanniana, muchos de estos textos. Con cierta ambición, se ha conceptualizado una base de datos donde se colijan los vínculos entre algunos de los testimonios y ediciones que forman parte de la tradición de muchas de estos documentos y da protagonismo a cada uno de ellos a través de registros individuales. Esto es entendible desde la perspectiva de la sociología de los textos, definida por McKenzie (2005), que entiende una obra a partir de sus múltiples materialidades conformadas a través del tiempo. En sintonía con esto, al hablar de fuentes no se habla solamente de las originales, sino de todas aquellas ediciones y versiones que a lo largo del tiempo se han ido produciendo, como veremos posteriormente.

A pesar de que la voluntad inicial fue la de crear una gran base de datos enlazada, con tablas interconectadas y elementos que reflejen y recojan la riqueza relacional entre los testimonios de las fuentes, fruto de décadas de investigación, se han debido tener en cuenta varios factores determinantes que responden a la realidad académica. No obstante, se debe destacar uno de ellos como especialmente relevante: es el de la autogestión que debe tener su mantenimiento.

Es de sobra conocido que los proyectos con un componente digital se deben enfrentar a una realidad compleja, en parte debido a su ambigua perdurabilidad fundamentada. Esto se debe a que la financiación de los proyectos está sujeta a la duración de los mismos y no siempre pueden ser continuados más allá del tiempo acotado. Ante esta inestabilidad aquellos que tienen una fuerte vertiente digital se encuentran con diversos obstáculos para proseguir. Existe un riesgo importante de que los frutos del proyecto puedan desaparecer al no existir un apoyo económico continuado ya que, por una parte, el uso de dominios o servidores conllevan un coste y, por otra, un proyecto de gran complejidad digital conlleva una dependencia de una estructura mínima de personal técnico informático que, en la actualidad, no puede ser asumida por la mayor parte de grupos y de centros de investigación del área de las Humanidades.<sup>13</sup>

Por este motivo, tal y como se ha comentado, la autogestión ha tenido que ser una característica prioritaria que ha formado parte en el diseño conceptual de la base de datos. Considerando que los que se encargan de recoger la información de las fuentes, incluirla en los registros y mantener viva la plataforma son profesionales de las Humanidades con conocimientos a nivel de usuario medio, sin necesidad de formación específica en competencias digitales, no conviene hacer complejas estas tareas en un estado inicial del desarrollo. Así pues, se debe simplificar al máximo tanto el uso como el mantenimiento del sistema. La meta a largo plazo, no obstante, no puede ser simplemente la de crear repositorio estático, sino que, más bien, se debe tener en consideración un progreso paulatino que aproveche los datos y metadatos que se vayan generando para continuar haciendo crecer la idea del proyecto, hasta llegar a un punto donde sea posible gestionar todo ello a través de sistemas más eficientes que procesen la información relacional de las fuentes.

---

<sup>13</sup> Para comprender una parte de las problemáticas vinculadas a las Humanidades Digitales, véase Grusin 2014. Si se quiere profundizar un poco más en el debate y como respuesta a dicho artículo, véase Adema y Hall 2016.

Al contemplar dichas especificidades y después de profundas reflexiones para conseguir dar respuesta tanto a los requisitos de calidad científica, como a la facilidad de uso de la plataforma para las personas que se encarguen de su mantenimiento, así como para los usuarios, se ha convenido que el desarrollo de una base de datos inicial debía hacerse en CSV. Este sencillo formato registra la información de una forma simple, abierta, fácilmente exportable y facilita una rápida adaptación a otros sistemas de gestión de bases de datos como podrían ser MySQL, FileMaker u Oracle Database, aunque con la prioridad siempre de utilizar *software* libre que se enfoque en la programación objeto-relacional para destacar las relaciones entre las fuentes históricas originales, sus materialidades y testimonios. La simplicidad del CSV permite que se aloje en casi cualquier sistema de gestión web, como por ejemplo Wordpress —que en el ámbito de la disposición de bancos de datos es más limitado—, además de agilizar la exportación de la información contenida en un documento CSV —como, por ejemplo, en hojas de cálculo Excel— y potenciar su reutilización.

Como punto de partida se ha tenido en cuenta el estudio y desarrollo de Hernández (2022: 431-441). En dicha investigación se parte de la creación de una serie de campos basados en el sistema Dublin Core, utilizado para la descripción de objetos digitales especialmente en «instituciones-memoria» y basado en XML, para crear un sistema de organización específico. Desde la creación de metadatos en Dublin Core, que adapta a las especificidades del proyecto, se establecen relaciones entre diversos elementos de diversas tipologías materiales —manuscritos, impresos, fotografías, postales...— relacionadas con el patrimonio cultural de la comedia *La dama boba* de Lope de Vega, que es el objetivo del estudio. Posteriormente hace uso del CSV para poder estructurar parte de esos metadatos y disponerlos en una página web que funciona como ejemplo expositor de algunos de los elementos recogidos. Además de esto, se centra en la creación de identificadores únicos de cada registro que no solo sean números progresivos, sino que los dota de significado para que por sí mismos indiquen la procedencia o tipología de los materiales.

La base de datos de VaHiFo, no requiere una complejidad tan grande en su desarrollo, ya que el objetivo es facilitar la información sobre las fuentes históricas valencianas, tanto de las originales —manuscritas e impresas— como ediciones de estas, no crear una exhibición directa de los materiales. Para su ejecución se parte del CSV, utilizado en la metodología anterior y que permite distribuir en líneas los registros sobre cada fuente histórica, pero los campos que comprenden están adaptados a las exigencias de los objetivos del proyecto y de las características de las fuentes. Se ha creado, por tanto, un sistema descriptivo general propio que permite registrar la información de una forma homogénea.

A continuación, se ofrece una descripción de cada uno de los campos que deben rellenarse individualmente para cada elemento, del tipo que sea. No obstante, a excepción del identificador único no todos los campos deben cumplimentarse, ya que la propia realidad material de los objetos lo impide —un manuscrito no tiene editorial, por ejemplo— tal y como se verá.

- *Identificador (Id.)*.

Este campo es único, obligatorio y clave para todos los registros de fuentes. Se forma por la conjunción de tres series de números que confieren singularidad y estructura a los elementos en la base de datos. El primer conjunto numérico se asocia con la obra original y abarca hasta tres cifras consecutivas, que comienzan en 001 y se pueden ampliar según sea necesario, y se aplican a todos los testimonios y ediciones vinculados a ella. La secuencia numérica de este conjunto depende exclusivamente del momento en que la obra original se incorpora a la base. Por ejemplo, los registros que comienzan con 001 se asignan a la Crónica de Jaume I.

Los siguientes dos dígitos, que comienzan por 00, determinan la secuencia de registro dentro de la tabla para cada elemento asociado a la misma obra principal. Estos números pueden formar un grupo específico dentro de la obra principal que refleje una tradición textual específica; por ejemplo, el código 00101... inicia el grupo correspondiente a la traducción latina de la Crónica de Jaume I.

Los dos dígitos finales indican el registro del elemento dentro de los grupos mencionados anteriormente. Comienzan por 00, que se utiliza principalmente para denotar la conceptualización de la obra original —en el caso de que vayan precedidos por otro 00— pero también, para indicar ediciones príncipes o primeros manuscritos. A partir de 01 se identifican las distintas materializaciones, ya sean copias o ediciones, que pertenecen a su grupo correspondiente.

De esta forma se obtiene un identificador único formado por la estructura nnn+nn+nn —donde «n» es el número. Según el ejemplo previo, 0010000 indicaría el concepto original de la Crónica de Jaume I —es decir, la idea platónica de obra y que, por tanto, no puede ser tangible—; 0010001 sería su primera edición crítica catalogada en la base —que con metodología ecdótica busca aproximarse lo máximo posible a la idea original de la obra—; y 0010100, el primer manuscrito de la versión en latín de Pere Marsili.

Este método de creación de identificadores tiene dos objetivos. Por una parte, que cada registro tenga un número que responda a su unicidad para procesos técnicos y, por otra, establecer mediante su creación una vinculación entre testimonios. De esta manera se refleja ya en la propia nomenclatura de cada registro un vínculo jerárquico propio y clasificado por obras. Es de crucial importancia que su creación sea rigurosa ya que es indispensable para la automatización de posibles procesos de conversión a otros sistemas de gestión de bases de datos más desarrollados que puedan incluir las relaciones entre las obras.

- *Título*.

Este campo contiene el nombre tradicionalmente asignado a la fuente original y por el cual es reconocida toda su tradición. Generalmente, este nombre coincide con el primer conjunto de dígitos del Id. En el caso de tener varios se separan por un punto y coma entre espacios. Algunos

ejemplos son: *Llibre dels fets*; *Crònica de Jaume I* —que indican un doble título—, *Dietari de Pere Joan Porcar* o *Crònica de Beuter*.

- *Título o rúbrica de testimonio o edición.*

Aquí se registra el nombre propio o por el que se encuentra catalogado el testimonio, edición de cualquier tipo, documento o texto. Como ejemplos derivados de los anteriores tenemos: *Crònica del rei en Jacme*, *Coses evengudes en la ciutat i regne de València* o *Primera Part de la Història de València*.

- *Autor/a.*

Campo designado para incluir el nombre de las personas a las que se le atribuye la autoría de la obra. Se indica siguiendo la estructura «Apellido, nombre». De esta forma: Jaume I; Porcar, Pere Joan o Beuter, Pere Antoni.

- *Fecha.*

Aquí se debe incluir la fecha de composición de los elementos registrados en la base de datos. El formato puede ser mediante un año específico, periodos entre años o, incluso, un siglo. Por ejemplo: XIII, 1585-1629 o 1995.

- *Género.*

Por género se entiende la categoría a la que pertenece la fuente registrada. Todos los elementos vinculados de una misma obra, es decir, sus testimonios y ediciones, mantendrán por norma general un género idéntico como son historiografía, memorialística, fuentes diplomáticas o poesía histórica, entre otros.

- *Tipología.*

En este campo se indica el tipo de objeto registrado. Estos pueden variar desde manuscritos o documentos diplomáticos hasta ediciones impresas de diversa índole: edición príncipe, edición crítica, edición facsímil, etc.

- *Responsable (copista, traductor, editor, etc.).*

En esta sección se incluye el nombre de cualquier tipo de responsable encargado de la elaboración del elemento registrado en el caso de que los haya. Puede indicarse, «copista» para copias manuscritas, «editores» para ediciones impresas o críticas, traductores, transcritores, etc.

- *Lengua.*

Como su nombre indica en este apartado se incluye el idioma en que está compuesto el objeto registrado. Por tanto, refleja cualquier lengua que se haya utilizado o se utilice en el territorio valenciano, así como algunas traducciones a otras distintas.

- *Ciudad.*

Ciudad de composición del ítem registrado. Suele referir al lugar de escritura en un manuscrito o documento diplomático o, en el ámbito editorial, al lugar de impresión.

- *Editorial.*

Campo que solo se utiliza en caso de libros y ediciones impresas para, como su nombre indica, referir a editorial o imprenta a cargo de su ejecución.

- *Notas.*

Aquí se puede indicar cualquier dato que se considere relevante respecto al elemento registrado. Es un campo de extensión abierta y que puede recoger información que varía desde referencias bibliográficas hasta estado del testimonio —si este es fragmentario, si tiene una digitalización parcial, etc. Este campo se utiliza especialmente para indicar el tipo de vinculación entre otros objetos registrados en la tabla a partir de su «Id», por ejemplo: *Còpia de 0010100* o *Es basa en 0010300*.

- *Institución (preservación).*

Este campo solamente se utiliza en determinados elementos registrados para indicar la institución donde se encuentra preservado el objeto físico al que hace referencia. Además, cuando el registro incluye un enlace al facsímil digital de dicho ítem, se indica que institución almacena dicha digitalización.

- *Signatura.*

Se recogen aquí las signaturas de cada registro por la que aparece catalogada la fuente de referencia en cada institución. En el caso de ejemplares reproducidos digitalmente se indica, si es posible, la signatura del objeto físico del que se ha extraído. Es especialmente importante en el caso de la mayoría de elementos almacenados en archivos, tales como documentos diplomáticos.

- *Disponibilidad digital.*

Mediante una opción binaria de «Sí» o «No» se indica si el elemento registrado contiene una reproducción digital disponible para consulta.

- *Enlace.*

En este campo se incluye el enlace al objeto digital que se almacena, generalmente, en repositorios externos para su consulta en el caso de exista disponibilidad digital.

## 5. Algunas particularidades

Como sucede con cualquier disposición en la que se establece una metodología unitaria para elementos de diversas tipologías, se pueden presentar algunas especificidades disruptivas que deben ser tenidas en cuenta. Esto es de suma importancia por dos motivos en especial: mantener un rigor científico que dote de valor académico a la base de datos y, además, intentar proseguir una normativa metódica que pueda ser interpretable por máquinas gracias a su regularidad. A continuación, se presentan algunos de los supuestos más frecuentes que se han detectado a la hora de añadir los registros al repositorio:

- *La tipología «obra» en diversos tipos de registros.*

Para hablar de este aspecto, se debe especificar que se parte del concepto más amplio de «obra» definido por Peter Robinson (2013) y su representación en el mundo digital. Siguiendo este concepto y tal como hemos visto en la descripción del campo «Id», se crea una agrupación en los registros de fuentes donde se identifica por separado la obra, sus testimonios y ediciones de cualquier tipología. Como se puede observar, el registro que acaba en «0000» se reserva para el concepto original de obra, cuya datación es el momento de concepción de la misma y no posee más título que el que se establece en dicho campo. Esto sucede, especialmente en el caso de grandes obras historiográficas, literarias o jurídicas, pero no así en el ámbito documental ni tampoco en el de trabajos que los editan, como veremos ahora. Todos los elementos acabados en «0000» pertenecen, por tanto, a la tipología material «obra».

Para los documentos diplomáticos hay que entender que, por su naturaleza, se generan, principalmente, no como obras ideadas, sino como fruto de la cotidianeidad y de labores específicas —sin entrar en el aspecto de intereses particulares o colectivos de su composición— y por ello no pueden existir más allá de su propio soporte físico.<sup>14</sup> Es por este motivo que en el caso de fuentes de esta tipología, no existe, en principio, un registro principal «obra» sino que el primero ya se especificará como elemento material. Un ejemplo de esto sería el registro 0310100, que corresponde a los documentos originales preservados en el Arxiu del Regne de València y editados en 2013 por José Manuel Iborra y Margarita Vila en 2013, *Cartes comercials i lletres de canvi de Francés Crespo, mercader valencià (1585-1601)*, con «Id.» 0310101 en la base de datos.

En el caso de las ediciones de investigaciones contemporáneas que trabajan conjuntos documentales, aunque se podrían considerar obras por sí mismas desde la perspectiva de la biblioteconomía, no podemos establecerlas al mismo nivel que las obras consideradas fuentes históricas originales. En este caso y para diferenciarlas, se designa un número propio del primer grupo de dígitos —nnn— pero no acaba en «0000», sino en «0001», porque se comienza a identificar desde el propio estudio sobre otras fuentes históricas. Por tanto, este tipo de estudios carece también de categoría «obra» como es el ejemplo del registro 0340001 perteneciente a *Vint processos criminals d'Albalat de la Ribera (1611-1666)*, volumen editado por Rafael Fresquet en 1997.

---

14 Para una mayor profundización en el tema, véase Galende y García (2003).

- *Volúmenes de obras y colecciones.*

Existe un doble criterio a la hora de registrar cuándo una obra posee varios volúmenes o partes. Por norma, si un trabajo tiene varios volúmenes se recoge como distintas materialidades suyas a través del conjunto numérico central del «Id» —nnn01nn, nnn02nn, nnn03nn...—, tal y como sucede por ejemplo con la *Crónica* de Pere Antoni Beuter. No obstante, cuando existe una enorme cantidad de testimonios de cada uno de los volúmenes individuales de la obra, hace que registrarlos como una sola obra sea insostenible. En este caso, conviene establecer una numeración individual, como si de obras independientes se tratara. No obstante, el campo «Título» corresponderá a la denominación general de todos los volúmenes y mantendrán el mismo en cada uno de los registros. Ejemplo de este supuesto son las crónicas de Rafael Martí de Viciano o de Gaspar Escolano.

En el caso de colecciones de investigación contemporáneas o series de publicaciones que bajo un mismo título edita distintas obras, no debe reflejarse como una sola agrupación—a lo sumo, se puede indicar en «Notas» si se considera conveniente. En esta situación, cada ejemplar de la colección se vincula a la fuente que edita. Quizá el paradigma más representativo de esto sea la colección *Fonts històriques valencianes* editado por el Servei de publicacions de la Universitat de València.

- *Documentos diplomáticos.*

En el apartado «La tipología “obra” en diversos tipos de registros», ya se ha especificado el porqué de que los documentos diplomáticos no tengan una tipología con la denominación «obra» y de cómo deben empezar a numerar su campo «Id». Sin embargo, hay más particularidades que por sus características, tienen que ser tenidas en cuenta.

De primeras, como es sabido, la documentación en archivos no suele tener un título propio, sino que se identifica por su signatura. Por este motivo, en el apartado correspondiente —Título o rúbrica de testimonio o edición— simplemente se indica la signatura de registro precedida por las iniciales de la institución. Esto no exceptúa que no haya que rellenar los campos «Institución» y «Signatura» aunque se duplique la información.

Debido a la ausencia de título, el campo homónimo tampoco debería ser rellenado. En cambio, si una publicación posterior que edite los documentos, le da un nombre unificado, su denominación será la que se imponga en el campo vacío «Título», pues esto contextualiza el grupo de elementos vinculados, sin perjuicio del formato de creación de «Id».

- *Elementos perdidos.*

Como es obvio, en el devenir del tiempo y por diversas causas que varían desde los problemas de preservación hasta por la injerencia humana, gran parte de los testimonios se han perdido. A veces, publicaciones antiguas o investigaciones ofrecen algo de información de alguno de estos ejemplares extraviados. Cuando esto sucede y existe documentación previa a su desaparición, puede ser

conveniente que estos queden registrados. Ejemplo de esto podría ser un manuscrito desaparecido, pero del que hubo constancia de su existencia. En el caso de considerarse su inclusión en la base de datos conviene indicarlo en el campo destinado a las anotaciones, así como alguna referencia bibliográfica previa que hable de su existencia.

- *Especificidades del software.*

Por las características técnicas del *software* utilizado para crear el CSV y del que se parte para su disposición en el portal web, los dígitos «0» a la izquierda del número identificador aparecen eliminados. Un ejemplo de ello se da ya en el del «Id» del *Llibre dels fets*, cuya nomenclatura 0010000 pasa directamente a 10000. Una posible solución en el futuro podría ser la de añadir un código alfabético al principio, por ejemplo, «vhf» delante del «Id» numérico y quedaría de la siguiente forma: vhf0010000. Este criterio reflejaría, además, su vinculación al proyecto VaHiFo.

## 6. Conclusiones

Como hemos visto a lo largo del artículo, el diseño de una base de datos y la conceptualización de las fuentes para su registro en ellas es algo complejo. No solamente hay que tener en cuenta los objetivos que se pretenden lograr en el proyecto, sino que también se debe prestar atención a necesidades técnicas, perdurabilidad en el futuro, posibilidad de continuación, sencillez de uso — tanto para quien rellena el repositorio como para quienes necesitan consultarlo— y una capacidad de reutilización de la información, como base de la ciencia abierta.

El planteamiento aquí ofrecido cumple a la perfección con los requisitos específicos que se han indicado en el apartado correspondiente. Esta formulación dota de una calidad académica razonable al banco de datos que desde un planteamiento simple tiene la capacidad de implementar actualizaciones en futuros proyectos, con garantía de perdurabilidad, evolución y autogestión, de forma que pueda servir de forma eficaz a la comunidad científica. En su conceptualización, se va más allá de una simple recopilación de datos, ya que, como valor añadido, se informa de las relaciones entre elementos que permiten dar relevancia a las investigaciones precedentes y, a su vez, se ofrece un panorama más visible de la tradición de las fuentes históricas valencianas. Todo esto, claro está, sin obviar uno de los aspectos más útiles: la localización tanto física como digital de la mayor parte de estas.

A pesar de todo, como todos los sistemas, no es perfecto y, por tanto, no está exento de cambios en el futuro. Algunos de los criterios aquí expuestos, como los correspondientes al quinto apartado, se han debido revisar constantemente para establecer una norma lo más metódica posible y ante la aparición de cualquier nueva particularidad se ha buscado una resolución eficiente para su registro. Todo esto con el objetivo adicional de establecer una regularidad que permita que el procesado de datos pueda ser interpretado por máquinas de forma automática en actualizaciones

venideras. Además, el diseño se ha simplificado lo necesario para permitir una adaptación a cualquier sistema de gestión.

A la hora de escribir este artículo se continúa añadiendo registros para conseguir aglutinar todos los que sean posibles. Nos encontramos, por tanto, ante un desarrollo aún vigente que responde a las necesidades de procesado del registro de las fuentes históricas valencianas, que posee una excelente proyección y que se puede revelar como una herramienta de gran utilidad para la comunidad científica.

Para concluir, como se ha mencionado, este no es un artículo que hable de los resultados del proyecto VaHiFo, sino solamente mostrar “el concepto de su sistema de información y servir de ejemplo para otros proyectos similares. Es por ello por lo que en este estudio de caso se explica todo el proceso de construcción, los desafíos de los que se parte y, por supuesto, la sistematización de los datos. Dichos elementos, en muchos casos, pueden ser cuestiones comunes a resolver en el ámbito de las Humanidades Digitales. De esta manera, se puede contemplar como los estudios de una revista especializada en estudios culturales y literarios son compatibles con perspectivas novedosas que influyen día a día en la evolución de estas áreas.

## Bibliografía

- Adema, J. y Hall, G. (2016) «Posthumanities: The Dark Side of ‘The Dark Side of the Digital’», *JEP. Disrupting the Humanities: Towards Posthumanities*, 19, 2. Disponible en: <https://quod.lib.umich.edu/j/jep/3336451.0019.201?view=text;rgn=main>.
- Agénjo Bullón, X. (2004) «De las bibliotecas virtuales a las ontologías y webs semánticas», en *II Congreso Nacional de Bibliotecas Públicas*, Madrid, Ministerio de Cultura, pp. 354-364. Disponible en: <http://hdl.handle.net/10421/788>.
- Ferrer Valls, T. (2018) «Hacia un sistema integrado de bases de datos sobre teatro clásico español y práctica escénica: el proyecto coordinado ASODAT», *Cuadernos AISPI*, 11, pp. 115-130. Disponible en: <https://www.ledijournals.com/ojs/index.php/cuadernos/article/view/1431>.
- Fiormonte, D.; Numerico, T. y Tomasi, F. (2015). *The Digital Humanist: A Critical Inquiry* (Desmond Schmidt, Christopher Ferguson trads.), Nueva York, Punctum Books. Disponible en: <https://www.jstor.org/stable/j.ctv1bd4hbj>.
- Galende Díaz, J. C. y García Ruipérez, M. (2003) «El concepto de documento desde una perspectiva interdisciplinar de la diplomática a la archivística», *Revista general de información y documentación*, 13, 2, pp. 7-35. Disponible en: <https://produccioncientifica.ucm.es/documentos/5d3999c32999520684452f10>.
- Grusin, R. (2014). «The Dark Side of Digital Humanities: Dispatches from Two Recent MLA Conventions», *Differences*, 25, 1, pp. 80-92.
- Hernández Tornero, C. (2022) *Investigaciones sobre la transmisión, recepción y tradición cultural de La dama boba: del manuscrito autógrafo a la era digital* [tesis doctoral], Universitat Autònoma de Barcelona.
- Jacinto, T. y Balsa de Pinho, J. (2008) «Instituciones de la Memoria. Reflexiones sobre sus particularidades», *Boletín de la ANABAD*, 58, 4, pp. 283-95. Disponible en: <https://www.anabad.org/wp-content/uploads/2011/03/2008.4.pdf>.
- López Poza, S. (2017) «La base de datos Symbola de divisas o empresas históricas. Planteamiento y diseño conceptual», *Studia Aurea*, 11, pp. 93-109. Disponible en: <https://pdfs.semanticscholar.org/992e/7e5b894348a8c3ad2288d311f283f6c5d158.pdf>.
- López Poza, S. y Pena Sueiro, N. (2020) «SYMBOLA, base de datos de divisas o empresas históricas. Un recurso sobre cultura visual, literatura e historia en internet», *Norba. Revista de Arte*, 40, pp. 129-155. Disponible en: <https://revista-norbaarte.unex.es/index.php/NRA/article/view/549>.
- López Yepes, J. (1998) «Las bases de datos históricas», *Anales de documentación*, 1, pp. 99-124. Disponible en: <https://dialnet.unirioja.es/servlet/articulo?codigo=241290&orden=0&info=link>.
- Malavassi Aguilar, A. P. (2012) «Las bases de datos como herramienta para la investigación histórica», *Diálogos: Revista electrónica de historia*, 13, 01, pp. 193-197. Disponible en:
- McKenzie, D. F. (2005) [1a ed. 1986] *Bibliografía y sociología de los textos*. Madrid: Akal.
- Petrucci, A. (1999) *Escritura de la memòria i memòries de l'escrit. De l'ordre dels objectes escrits al desordre de l'escritura virtual*. García Porcar, R. M<sup>a</sup> Gregori Roig y R. J. Puchades Bataller (trads.), Valencia, Universitat de València.
- Robinson, P. (2013) «The Concept of the Work in the Digital Age». *Ecdotica*, 10, pp. 13-42.