

Recuperando información histórica en astronomía

Un viaje al pasado

para ayudar a conocer el futuro



FIGURA 1 Algunos documentos astronómicos históricos: (A) Carta celeste pictórica (Hevelius, 1690). (B) Carta celeste no pictórica (Cottam, 1889). (C) Globo celeste (Mercator, 1551). (D) Planisferio-catálogo (Doppelmayr, 1742). (E) Carta cosmológica (Doppelmayr, 1742). (F) Tabla principal de un catálogo celeste (Gill, 1898). (G) Carta fotográfica del cielo (ROA, 1923). (Excepto donde se indique, todas las imágenes son cortesía de los autores)

Archivos y bibliotecas astronómicas de todo el mundo albergan miles de documentos históricos en soportes no electrónicos, los cuales contienen valiosa información a la que es muy difícil acceder de forma selectiva, debido a la insuficiente descripción documental que suele realizarse del contenido de estos recursos en las bases de datos de estas instituciones.

FRANCISCO JAVIER MOLERO MADRID Y MARÍA PILAR ALONSO LIFANTE

El aumento de las posibilidades de acceso a la información a través de las Tecnologías de la Información en general, y de Internet en particular, han acelerado el proceso de migración de la conocida como «Sociedad de la Información» a la «Sociedad del Conocimiento». El incremento exponencial de todo tipo de información que se vierte diariamente a la red está planteando retos, hasta hace unas décadas casi inimaginables, en lo que se refiere a la clasificación, catalogación, almacenamiento y recuperación de dicha información, lo que hoy conocemos como «Web Semántica».

Estamos acostumbrados a leer en numerosas revistas y sitios web especializados, el enorme volumen de información que generan proyectos de observación y misiones espaciales. Sirva como referencia los 50 Gb diarios de datos que se anuncia que proporcionará el observatorio GAIA de la Agencia Espacial Europea durante sus cinco años de misión. La existencia de esta ingente cantidad de información y su falta de estandarización, ha propiciado la proliferación de iniciativas encaminadas a unificar el modo en que se almacena, analiza y difunde esta información, cu-

yo máximo exponente es la «Alianza Internacional de Observatorios Virtuales» (IVOA, por sus siglas en inglés). Entre los muchos objetivos que persiguen las instituciones pertenecientes a IVOA, está el favorecer la interoperabilidad de los datos a través de la creación de estándares que permitan el intercambio de información de forma unificada. El equipo del «Observatorio Virtual Español» (SVO) se encuentra ubicado en el Centro de Astrobiología (INTA-CSIC) en Madrid.

La información astronómica puede clasificarse en dos grandes grupos atendiendo al soporte en el cual esta viene dada. Por un lado disponemos de datos en formato electrónico, es decir, toda aquella información que se obtiene de forma automatizada gracias a mo-

En la astronomía se ha pasado de la búsqueda de «documentos» a la búsqueda de «datos»

ernos dispositivos (nuevos telescopios, cámaras CCD, software informático, etc.) y que se almacena en bases de datos accesibles solo electrónicamente. Sin embargo, existe también un vasto volumen de información histórica almacenada en dis-

tintos soportes no electrónicos (placas fotográficas, atlas y catálogos de estrellas en formato libro, mapas celestes, imágenes astrofotográficas, ver Figura 1), que se conservan en archivos, cartotecas y bibliotecas astronómicas de todo el mundo, y que todavía son de gran interés para la comunidad científica. En efecto, grandes descubrimientos en astronomía han sido posibles gracias a la comparación de datos históricos y contemporáneos. Sirva de ejemplo el descubrimiento del ahora bien conocido movimiento propio de las estrellas por parte del astrónomo inglés Edmund Halley, cuando comparó las posiciones de estas estrellas en observaciones realizadas en su época (siglo XVIII) con aquéllas efectuadas por el griego Hiparco y otros en el siglo II.

La información almacenada de forma electrónica ha dado lugar a la aparición de grandes bases de datos especializadas en astronomía como SIMBAD (Set of Identifications Measurements and Bibliography for Astronomical Data) del Centro de Datos Astronómicos de Estrasburgo (CDS), NED (NASA/IPAC Extragalactic Database), o la base de datos de catálogos Vizier, también del CDS. El amplísimo abanico de op-

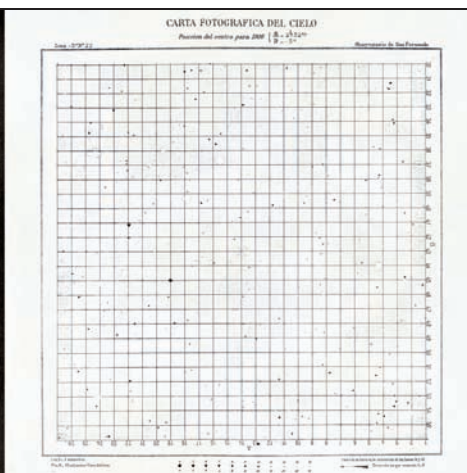


FIGURA 4 A la izquierda, impresión fotográfica de un dibujo de la nebulosa de Orión. A la derecha, una carta fotográfica del cielo correspondiente al proyecto «Carte du Ciel» [ROA, 1923].

mejora de la descripción del contenido de cualquier recurso en general, y de los astronómicos en particular, pasa en muchos casos por la creación y utilización de nuevos metadatos que permitan una recuperación de información verdaderamente satisfactoria en los OPAC de archivos, bibliotecas y cartotecas.

POSIBLES SOLUCIONES

Existen soluciones de diversa índole al problema que se acaba de plantear. La primera de ellas consiste en digitalizar con reconocimiento óptico de caracteres todos los documentos de los que se dispone. De esta forma, no solo se podría buscar en el interior de los documentos las cadenas de caracteres por los que cualquier usuario decidiera buscar un determinado recurso, sino que este usuario también podría ac-

lo que conlleva un coste económico difícil de asumir en estos momentos. Este es precisamente uno de los motivos por los que solo un pequeño porcentaje de atlas y catálogos de estrellas históricos se encuentran completamente digitalizados.

Este obstáculo económico, agudizado por la actual situación de crisis financiera, abre la puerta al estudio de nuevos metadatos descriptivos que permitan realizar registros bibliográficos más completos de los documentos conservados. Así, aunque no se pueda acceder a los datos originales digitalizados, podremos consultar los OPAC y recuperar aquellos documentos que contienen exactamente los datos deseados. En este sentido, hemos realizado dos análisis sobre atlas [2] (Figura 2) y catálogos de estrellas [3] (Figura 3) de los siglos XVIII, XIX y XX procedentes del Real Instituto y

Observatorio de la Armada (ROA) (San Fernando, Cádiz) con objeto de obtener los parámetros técnicos mínimos deseables que sería conveniente registrar de estos dos conocidos documentos astronómicos. También puede consultarse otro estudio

sobre objetos celestes presentes en imágenes astrofotográficas históricas [4] (Figura 4).

Los resultados [5] muestran, por un lado, que existe un cierto número de estos parámetros técnicos que no están presentes en los estándares de catalogación y que deberían tenerse en cuenta en las descripciones bibliográficas, y por otro, que hay in-

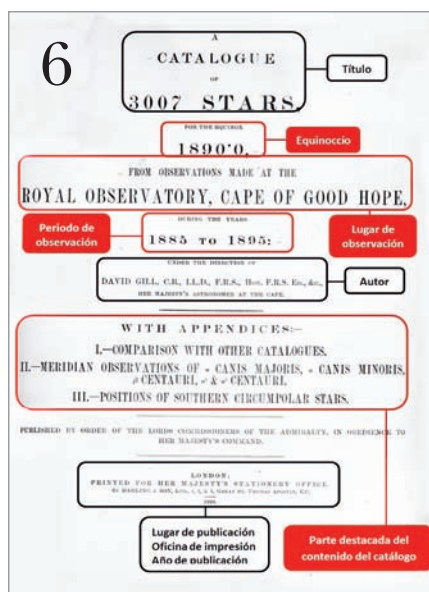
formación que puede ser descrita actualmente con dichos estándares, pero cuyos campos no están siendo utilizados ampliamente por parte de los catalogadores. Nuestro análisis y propuesta de nuevos campos para los estándares de catalogación, supone un paso adelante en el complejo proceso de descripción de recursos altamente especializados, lo cual permitirá a sus potenciales usuarios poder buscar en los OPAC de bibliotecas de una forma más próxima a como lo hacen en otras bases de datos astronómicas especializadas. Creemos, además, que incorporar todos estos parámetros a los estándares de catalogación supondría un primer paso en el proceso paulatino de migración que debe producirse de dichos estándares generales a los desarrollados por los observatorios virtuales.

METODOLOGÍA

La mejora de la descripción documental de cualquier recurso o documento, pasa inevitablemente por el profundo conocimiento del contenido de los mismos, lo cual requiere analizarlos cuidadosamente en busca de la presencia de aquellos datos de mayor interés para los investigadores. De entre los distintos lugares de España donde poder abordar un estudio de estas características, la biblioteca del ROA es uno de los centros donde se conserva mayor volumen de información astronómica histórica (ver *Astronomía* 179, 30-37). En concreto, se analizaron un total de 22 atlas y 28 catálogos de estrellas de distintos países que abar-

Existe un vasto volumen de información histórica almacenada en bibliotecas astronómicas de gran interés para la comunidad científica

ceder al propio documento digitalizado, sin necesidad de desplazarse físicamente al lugar donde se conserva el ejemplar original. Sin embargo, teniendo en cuenta que existen instituciones que pueden llegar a conservar decenas de miles de documentos con centenares de páginas cada uno, estaríamos hablando de millones de páginas a digitalizar,





RED DE BIBLIOTECAS DE DEFENSA

7

Título: A catalogue of 3007 stars, for the equinox 1890.0, from observations made at the Royal Observatory, Cape of Good Hope during the years 1885 to 1895 / David Gill

Editorial: London : Darling & Son, 1898

Descripción física: XLIV, 263 p. : 33 cm

Materias: [Estrellas](#)
[Catálogos](#)
[Astronomía](#)
[Gill, David](#)

Title: A Catalogue of 3007 Stars, for the Equinox 1890.0, from observations made at the Royal Observatory, Cape of Good Hope, during the years 1885 to 1895; under the direction of David Gill, etc.

Author: Royal Observatory, Cape of Good Hope

Contributor: David GILL Sir, K.C.B.

Publication Details: London, 1898.

Identifier: System number 000599493

Physical Description: xliv, 263 p. : 4°.

Shelfmark(s): General Reference Collection 8563.g.14.
Document Supply OP-GPA/16790

UIN: BLL01000599493

Corporate name: [Royal Observatory, Cape of Good Hope](#)

Main title: A catalogue of 3007 stars, for the equinox 1890.0, from observations made at the Royal observatory, Cape of Good Hope, during the years 1885 to 1895; under the direction of David Gill ... With appendices:--I.--Comparison with other catalogues. II.--Meridian observations of [alpha] Canis majoris, [alpha] Canis minoris, [beta] Centauri, [alpha] * & [alpha] * Centauri. III.--Positions of southern circumpolar stars. Published by order of the lords commissioners of the Admiralty, in obedience to Her Majesty's command.

Published/Created: London, Printed for H.M. Stationery off., by Darling & son, Ltd., 1898.

Description: xliv, 263 p. tables. 33 x 26 cm.

LC copy: QB6 .C2 1890 Copy 2 Copy 2.

LCCN: 01005452

Subjects: [Stars--Catalogs](#).

Notes: Binder's lettering: Cape catalogue 1885-1895.

LC classification (partial): [QB6](#)

Related names: [Gill, David, Sir, 1843-1914](#)
[Great Britain. Admiralty](#)

Type of material: Book

FIGURA 5 Imagen de uno de los autores de este trabajo junto al Dr. Francisco José González González, Director Técnico de la Biblioteca y el Archivo Histórico del ROA, justo en el lugar donde se conservan las placas fotográficas del proyecto «Carte du Ciel». **FIGURA 6** Información común (recuadros blancos) y técnica (recuadros rojos) que ofrece la portada del catálogo de estrellas de Gill (1898). **FIGURA 7** Registro del catálogo de estrellas de David Gill creado por el ROA, BL y LC. Consulta: febrero 2014. (Bibliodef, British Library y Library of Congress)

caban el periodo histórico comprendido por los siglos XVIII, XIX y XX. El objetivo, además de extraer los datos de mayor interés, consistía en tratar de localizar patrones de información en la estructura organizativa de estos documentos, es decir, com-

probar si determinado tipo de información solía ofrecerse más frecuentemente en determinadas secciones de esos atlas y catálogos. De esta forma, una vez conocidos los datos que deben considerarse en los registros bibliográficos, es posible facilitar

la labor de los catalogadores, indicándoles la probabilidad que existe de localizar esos datos en las distintas secciones de los atlas y catálogos, agilizando así el proceso de catalogación de unos documentos que suelen ser bastante especializados y extensos.

En concreto, la localización de los documentos a analizar se realizó mediante consultas genéricas al Catálogo Colectivo de la Red de Bibliotecas de Defensa (comúnmente conocido como Bibliodef). Tras la selección de la muestra, con respecto a los atlas se llevó a cabo un estudio preliminar sobre su contenido, analizando portadas (Figura 6), índices, introducciones, cartas de estrellas, etc. Esto nos permitió desarrollar un nuevo listado con los datos proporcionados con mayor frecuencia. A raíz de este estudio, se localizó un cierto número de parámetros, y más tarde se llevó a cabo una revisión atlas por atlas para conocer la frecuencia de aparición de cada parámetro en cada sección sobre el total de la muestra. Con respecto a los catálogos, el método fue muy similar. Se llevó a cabo un primer análisis considerando solo los datos proporcionados por la «tabla principal» del catálogo, lo que permitió identificar los nombres de las columnas que conforman las tablas de datos de los mismos, y que son los datos principales a registrar que aportan dichos catálogos. El resto de secciones fueron analizadas siguiendo los mismos criterios que en el caso de los atlas.

SOBRE LA PROPUESTA DE PARÁMETROS TÉCNICOS

Como resultado del estudio que se acaba de resumir, se ha extraído un conjunto de parámetros de descripción que puede clasificarse en distintos grupos, bien atendiendo a la naturaleza de los mismos (Tabla 1), o bien a su encaje en los actuales estándares de catalogación y codificación (Tabla 2). Según su naturaleza, estos parámetros pueden ser de tipo astronómico, cartográfico o documental. Según su encaje en los estándares,

estos parámetros pueden clasificarse a su vez en tres grupos: (1) parámetros que ya se encuentran presentes en los estándares empleados en archivos y bibliotecas; (2) nuevos parámetros que podrían ser registrados en algunos de los campos de dichas normativas, aunque estos no hayan sido originalmente diseñados para registrar estos nuevos parámetros; y (3) nuevos parámetros que solo podrían tener cabida en campos generales de notas, y que serían susceptibles de convertirse en nuevos campos de los citados estándares.

Con respecto a esta última clasificación, es necesario aclarar que las normas de catalogación han creado lo que se conoce como «campos de notas». Se trata de un *cajón de sastre* donde los catalogadores pueden indicar en los registros bibliográficos aquella información relevante que creen que puede ser de utilidad para localizar el documento, pero que no tiene cabida en otros campos de los estándares. En cambio, lamentablemente todavía existen sistemas de recuperación de información que no tienen en cuenta las informaciones anotadas en campos de notas, debido principalmente a dos razones. La primera es que la mayoría de anotaciones que se realizan en estos campos son en lenguaje natural, y no están estandarizadas, con la dificultad que ello conlleva a la hora de encontrar coincidencias entre la información registrada y la cadena de caracteres empleada por los usuarios para hacer sus búsquedas. Y en segundo lugar, si se suprimen las búsquedas en determinadas partes del registro bibliográfico, se gana en velocidad optimizando el tiempo de búsqueda aunque en detrimento de la exhaustividad de la misma.

EJEMPLO PRÁCTICO DE UNA CONSULTA A UN OPAC

Supongamos que un investigador accede al OPAC de una biblioteca astronómica y desea consultar todos los catálogos de estrellas del siglo XIX de esa institución que proporcionan explícitamente movi-

TABLA 1

Tipos de parámetros	Parámetros técnicos identificados
Parámetros cartográficos	Coordenadas; tipo de coordenadas; escala; perspectiva; proyección.
Parámetros astronómicos	Época; equinoccio; magnitud; tipo de magnitud; instrumentos; hemisferio celeste; lugar de observación; periodo de observación; nombre de las constelaciones; nomenclatura de las estrellas; imágenes de constelaciones; longitud de onda; objetos de interés; significado de las columnas de la tabla principal del catálogo (movimiento propio, precesión, correcciones, etc.); título de otras tablas que complementan la tabla principal del catálogo.
Parámetros documentales	Documento relacionado; índice; tipo de documento.

Clasificación de parámetros para la descripción de catálogos y atlas de estrellas históricas atendiendo a su naturaleza.

TABLA 2

Parámetros que pueden describirse con los estándares de catalogación Coordenadas; época/equinoccio; proyección; escala.
Nuevos parámetros que podrían registrarse en los actuales campos de descripción Nombre de las constelaciones; hemisferio celeste; índice; magnitud; periodo de observación; lugar de observación; documento relacionado.
Nuevos parámetros que podrían registrarse solo en campos de notas Instrumentos; significado de las columnas de la tabla principal de catálogo; perspectiva; nomenclatura de las estrellas; tipo de coordenadas; tipo de magnitud; longitud de onda; objetos de interés; leyenda de otras tablas que proporcionan información extra sobre correcciones, instrumentos, etc.; imágenes de constelaciones.

Clasificación de parámetros para la descripción de catálogos y atlas de estrellas históricas atendiendo a su encaje en los actuales estándares de catalogación.

mientos propios de las mismas. Esta búsqueda, aparentemente sencilla para una base de datos astronómica tipo Vizier, resulta muy complicada a través del OPAC. Para comprobarlo, tomaremos como ejemplo el catálogo de estrellas de David Gill de 1898 que se conserva tanto en el ROA como en la British Library (BL) y la Library of Congress (LC), con objeto de comparar sus registros bibliográficos. La Figura 7 muestra cómo se ha catalogado la

obra por parte de las tres bibliotecas mencionadas.

Como puede observarse, el registro creado por el ROA solo ofrece al usuario el título del documento y algunos datos de publicación. Sin embargo, consultando la obra original (véase de nuevo la Figura 3), puede verificarse, por un lado, que la tabla principal de este catálogo facilita una columna con el movimiento propio anual de las estrellas observadas y, por otro, que la pro-

pia portada del mismo ofrece cierta información astronómica de interés como el equinoccio (1890.0, que las normas actualmente permiten describir pero no se indica en este caso), el lugar de observación (Real Observatorio del Cabo de Buena Esperanza) o el periodo de observación (del año 1885 a 1895). Nótese que esta información se ha indicado en la etiqueta «Título» del registro, pero no se ha reflejado en sus campos propios (existentes en la mayoría de los estándares), lo cual permitiría crear diferentes puntos de acceso a la información y repercutiría en una mejor recuperación de información. Sucede lo mismo en los registros que ofrecen BL y LC. Además, aunque las tres bibliotecas han empleado el formato MARC como herramienta de codificación, ni siquiera hay unanimidad en la presentación de la información que se

Los campos de notas son un cajón de sastre de los estándares donde los catalogadores pueden indicar otras informaciones relevantes

muestra al usuario. En cualquier caso, ninguno de los tres registros indica que el catálogo incluye información sobre movimientos propios de las estrellas.

En definitiva, si un investigador desea realizar una búsqueda de todos los catálogos de estrellas del siglo XIX que proporcionan explícitamente el movimiento propio, no recuperará la totalidad de documentos relevantes que conserva la biblioteca, ya que, en muchos casos, esa información no ha sido incorporada a los registros bibliográficos y por ende a la base de datos.

CONCLUSIONES

Podemos concluir que son dos las principales dificultades que encuentra un investigador en astronomía cuando accede al OPAC de la biblioteca astronómica. Por un

lado, la interfaz de búsqueda no se parece demasiado a la de las bases de datos especializadas en las que está acostumbrado a consultar y, por otro, cuando realiza una búsqueda por parámetros astronómicos/astrofísicos puede recuperar un número muy reducido de documentos (o ninguno en muchas ocasiones) lo cual no refleja en absoluto la cantidad ingente de información histórica que alberga la biblioteca. Esto indica que los estándares de catalogación no están preparados para abordar una adecuada descripción de los recursos astronómicos históricos, a pesar de los grandes esfuerzos que se han venido realizando desde las agencias de normalización en los últimos años. Por tanto, esta propuesta de parámetros de descripción se presenta como una alternativa real para satisfacer las necesidades

de información histórica de los investigadores, aunque creemos que la solución a parte de los problemas comentados podría alcanzarse de la mano de los conocidos «esquemas de metadatos descriptivos» (Dublín Core, MARCXML, MODS, etc.), gracias a que estos permiten

la adición de parámetros y la creación perfiles de aplicación específicos, algo que es más difícil de conseguir con los clásicos estándares de catalogación.

Finalmente, es importante señalar que los parámetros que proponemos en nuestro trabajo para mejorar la descripción bibliográfica de recursos astronómicos históricos, van a ser también propuestos para su incorporación en la International Standard Bibliographic Description (ISBD), uno de los estándares de catalogación internacionales más empleados en todo el mundo. Además, la Biblioteca Nacional está dando los primeros pasos para organizar cursos de formación de su personal, con objeto de que se familiaricen con los recursos astronómicos históricos así como con el renovado proceso de

descripción documental que proponemos llevar a cabo de dichos recursos. (A)

BIBLIOGRAFÍA

- [1] Alonso Lifante, M. P. et al. «La recuperación de información en astronomía: una propuesta para mejorar la descripción documental de catálogos y atlas de estrellas históricos» en: J. A. Garzón-Guerrero y A. R. López-Sánchez (eds.), Libro de Actas del XXI Congreso Estatal de Astronomía (pp. 265-272), Granada, Red Andaluza de Astronomía, 2014. Disponible en: media.wix.com/ugd/b730a9_1e7501d2b14f48aca5404fe9cb8ace56.pdf
- [2] Alonso Lifante, M. P. et al. «Historical celestial cartography: a proposal to improve the documental description of the contents of star charts and atlases». *The Cartographic Journal*, published online 31/08/2014 [advanced articles]. Disponible en: dx.doi.org/10.1179/1743277414Y.0000000093.
- [3] Alonso Lifante, M. P. et al. «A proposal to improve the description of astronomical resources: the case of historical star catalogues». *Journal of Documentation*, 71 (2) [en prensa].
- [4] Alonso Lifante, M. P. et al. «Google Sky y los estándares de catalogación: un ejemplo de divergencia entre la información astronómica existente y la que se puede describir». *Revista Española de Documentación Científica*, 36(4), 1-17, 2013. Disponible en: dx.doi.org/10.3989/redc.2013.4.1000
- [5] Alonso Lifante, M. P. *Recuperación de información astronómica: nuevas perspectivas de análisis de la cartografía celeste histórica*. Tesis Doctoral, Universidad de Murcia, 2014. Disponible en: digitum.um.es/xmlui/bitstream/10201/42086/1/Tesis-M%C2%AA PilarAlonsoLifante.pdf

Francisco Javier Molero Madrid, Ingeniero Superior de Telecomunicaciones. Doctor en Ciencias, especialidad en Astrodinámica (Universidad de Murcia).



María Pilar Alonso Lifante, Licenciada en Documentación. Doctora en Ciencias Sociales, especialidad en recuperación de información astronómica histórica (Universidad de Murcia).

