HILOS DOCUMENTALES / Núm. 10, 2026 / ISSN 2618-4486
https://revistas.unlp.edu.ar/hilosdocumentales / archivohistorico@presi.unlp.edu.ar
RED DE ARCHIVOS DE LA UNIVERSIDAD NACIONAL DE LA PLATA, ARGENTINA

MODELADO DE DATOS EN WIKIDATA
UNA PROPUESTA PARA UNIFORMIZAR LA DESCRIPCIÓN DE AUTORIDADES DE PERSONAS

DATA MODELING IN WIKIDATA
A PROPOSAL TO STANDARDIZE THE DESCRIPTION OF PERSONAL AUTHORITIES

Cláudia De Souza*
Universidad de Puerto Rico Recinto de Río Piedras, Puerto Rico
claudia.souza@upr.edu
https://orcid.org/0000-0002-4168-9399

Dinah M. Wilson-Fraites**
Universidad de Puerto Rico Recinto de Río Piedras, Puerto Rico
dinah.wilson@upr.edu
https://orcid.org/0009-0009-4762-1625

Disponible en doi: https://doi.org/10.24215/26184486e080

Fecha de envío: 30 de abril de 2026
Fecha de aceptación: 10 de julio 2026
Fecha de publicación: 22 de septiembre de 2026

Esta obra está bajo una Licencia Creative Commons Atribución -NoComercial -CompartirIgual 4.0 Internacional

*Doctora en Documentación por la Universidad Carlos III de Madrid, España (2018). Actualmente es profesora catedrática asociada en la Universidad de Puerto Rico, Recinto de Río Piedras, donde imparte cursos en el Programa Graduado en Ciencias de la Información. Su docencia se enfoca en la organización y recuperación de la información, con especial atención al análisis, evaluación y diseño de bibliotecas y archivos digitales. Su principal línea de investigación se articula en torno a la ciencia abierta entendida como un bien común. Sus intereses incluyen aumentar la visibilidad, interoperabilidad y accesibilidad de los recursos de información relacionados a Puerto Rico.

**Bibliotecaria de catalogación en el Sistema de Bibliotecas de la Universidad de Puerto Rico, Recinto de Río Piedras. Posee una Maestría en Ciencias de la Información (2018) y una Maestría en Literatura Comparada (2011).

RESUMEN

Wikidata, la base de datos libre y abierta de la Fundación Wikimedia, puede ser consultada y editada tanto por personas como por máquinas, ya que organiza los metadatos de manera estructurada mediante el modelo RDF (*Resource Description Framework*). Esto facilita la integración de los recursos, la interoperabilidad y la visibilidad de la información en la web a nivel mundial. En los últimos años, su uso se ha extendido en galerías, bibliotecas, archivos y museos, donde se enlazan con otras bases de datos, repositorios y buscadores en internet, promoviendo el acceso abierto a los datos. Para garantizar estos beneficios, es fundamental que Wikidata se construya sobre modelos y esquemas de representación consistentes, principalmente porque son creados y gestionados de forma colaborativa por una comunidad de voluntarios. No obstante, la cantidad y diversidad de información generada dificulta mantener la precisión y uniformidad en los registros, especialmente en la descripción de individuos, cuyos datos pueden presentarse de formas distintas. En respuesta a esta problemática, el presente trabajo propone un modelo de propiedades en Wikidata para datos de autoridades de personas, orientado a estandarizar su descripción, mejorar la consistencia de los elementos y potenciar su accesibilidad y reutilización en un contexto de pensar globalmente y actuar localmente.

PALABRAS CLAVE

Wikidata, datos abiertos enlazados, normalización de registros, control de autoridades, interoperabilidad.

ABSTRACT

Wikidata, the free and open database of the Wikimedia Foundation, can be queried and edited by both people and machines, as it organizes metadata in a structured manner using the RDF (Resource Description Framework) model. This facilitates resource integration, interoperability, and the visibility of information on the web at a global level. In recent years, its use has expanded in galleries, libraries, archives, and museums, where it is linked to other databases, repositories, and internet search engines, promoting open access to data. To ensure these benefits, it is essential that Wikidata be built on consistent models and representation schemes, mainly because it is created and managed collaboratively by a community of volunteers. However, the amount and diversity of generated information make it difficult to maintain accuracy and uniformity in records, especially in the description of individuals, whose data may be presented in different ways. In response to this issue, the study proposes a property model in Wikidata for authority data on persons, aimed at standardizing their description, improving the consistency of items, and enhancing their accessibility and reuse in a context of thinking globally and acting locally.

KEYWORDS

Wikidata, linked open data, record normalization, authority control, interoperability

Introducción

En la era digital, los metadatos organizados de manera estructurada han adquirido gran relevancia, lo que ha impulsado el desarrollo de mecanismos que permiten que la información se comparta y se reutilice entre distintas plataformas, instituciones y comunidades. Actualmente, muchas de estas iniciativas se enfocan en los datos abiertos enlazados, con el objetivo de facilitar su procesamiento automatizado. Una de las más relevantes es Wikidata, el proyecto hermano de Wikipedia. Se trata de una base de datos abierta y multidisciplinaria, creada por la Fundación Wikimedia, que, desde su lanzamiento a finales de 2012, ha crecido de manera exponencial en términos de contenido, convirtiéndose en una referencia internacional con aproximadamente 140 millones de elementos registrados actualmente.

El contenido de Wikidata está disponible bajo una licencia libre, lo que permite que pueda ser consultado, reutilizado y vinculado a otros conjuntos de datos en la web. Esto se realiza mediante propiedades que describen las relaciones entre los elementos, organizadas en forma de sentencias (conocidas como declaraciones) siguiendo el esquema de tripletas sujeto-predicado-objeto del formato RDF (*Resource Description Framework*). Gracias a ese modelo flexible, Wikidata permite describir y enlazar prácticamente cualquier tipo de entidad: conceptos, objetos, eventos, organizaciones e incluso personas, quienes pueden ser representadas como autoridades. Este trabajo se realiza principalmente gracias a la labor de miles de voluntarios alrededor de todo el mundo, quienes contribuyen a agregar, revisar y mantener la información.

Como base multilingüe, Wikidata puede proporcionar datos en cualquier contexto e idioma, siempre que estén disponibles. Numerosos estudios la han explorado desde diferentes perspectivas, desde su estructura interna (incluyendo tanto los datos como la comunidad), hasta la calidad de los datos, analizando su completitud y cobertura. Asimismo, se ha abordado la dimensión de la ingeniería, evaluando las herramientas necesarias (Farda-Sarbas & Müller-Birn, 2019); y la aplicación práctica, mediante estudios de caso sobre el uso de Wikidata en proyectos de diversas disciplinas, desde las Ciencias de la Vida (Waagmeester et al., 2020) hasta la investigación en Ciencias Sociales (Arroyo-Machado, 2024).

Estos estudios muestran el potencial de Wikidata, pero también evidencian que el éxito de la plataforma depende no solo de su estructura y herramientas, sino de cómo los usuarios crean y gestionan la información (Piscopo, Phethean, & Simperl, 2017). En este sentido, el cumplimiento de las normas y restricciones establecidas por la comunidad resulta fundamental para evitar inconsistencias y preservar la calidad de la información. No obstante,

Pastor-Sánchez (2021) ha resaltado la heterogeneidad de criterios a la hora de utilizar una u otra propiedad para jerarquizar los ítems. La ausencia de un modelo de metadatos estandarizado genera inconsistencias en las descripciones de autoridades, dificultando la recuperación de información precisa, provocando vacíos en los datos y limitando la interoperabilidad entre distintos repositorios y sistemas de información. Esto puede derivar en errores en la identificación de personas u organizaciones, duplicación de registros y menor confiabilidad de los datos abiertos.

Algunos otros autores también reflejan esta inquietud. Por ejemplo, según Brasileiro et al. (2016), existen problemas en las jerarquías taxonómicas de Wikidata, lo que genera patrones de error recurrentes debido al uso incorrecto de ciertas propiedades. Frente a esto, estudios recientes han abordado la problemática mediante el uso de Wikidata como plataforma para estandarizar la descripción de autoridades, creando modelos de datos que permiten vincular registros con identificadores externos como el Virtual International Authority File (VIAF), Open Researcher and Contributor ID (ORCID) o *Library of Congress Name Authority File* (LCNAF). Un ejemplo destacado es el proyecto (*In)Visible Women* del Smithsonian, que recopiló más de 20 mil registros de personas en distintas unidades institucionales, muchos de ellos sin identificadores de autoridad, y desarrolló un modelo colaborativo en Wikidata para unificar y desambiguar esta información (Sorensen et al., 2026).

De acuerdo con Agenjo-Bullón y Hernández-Carrascal (2018), merece especial atención uno de los acontecimientos más significativos en la biblioteconomía mundial de los últimos años: la extensión del control de autoridades al ámbito de Wikipedia. Dicho proceso ha permitido vincular grandes volúmenes de datos de los ficheros de autoridades de las instituciones culturales con Wikipedia, dando lugar a la creación de Wikidata. En este sentido, Wikidata forma parte de iniciativas Open GLAM (del inglés *Open Galleries, Libraries, Archives and Museums*), que promueven la apertura, el acceso y la reutilización de la información del patrimonio cultural a través de plataformas colaborativas y los datos abiertos enlazados.

Bajo este contexto, surge la necesidad de establecer un marco claro de normalización, para garantizar que los metadatos se gestionen de manera consistente y confiable, asegurando la interoperabilidad entre distintos sistemas de información. Diversas comunidades han desarrollado esquemas de metadatos específicos para sus necesidades (Vellucci, 2004) y, a lo largo del tiempo, este control ha pasado de un enfoque estrictamente organizativo a uno que también cumple un propósito informativo (Hernández Sánchez, 2022). Los datos abiertos enlazados permiten la interconexión de bases de datos, aportando descripciones de metadatos

que, en su integración, permiten inferir conocimiento. De ahí la importancia de la estandarización de la descripción de conjuntos de datos y la conveniencia de la normalización de identificadores, lenguajes descriptivos, ontologías y vocabularios. De no evolucionar en ese camino, la no normalización puede impedir una interoperabilidad consistente.

Indudablemente, desde el punto de vista de la satisfacción del usuario, la posibilidad de enriquecer y estandarizar semánticamente los metadatos mejora su usabilidad, su consulta y abre numerosas vías de navegación. Por ello, la vinculación con los vocabularios de valores disponibles en un dominio concreto no debe ser una posibilidad más o menos realizable, sino una técnica que debe incorporarse a las tareas de control de autoridades (Agenjo-Bullón & Hernández-Carrascal, 2018). Esto no solo permite organizar y estandarizar los nombres para garantizar consistencia y precisión en los registros, sino que también asegura la identificación unívoca de personas, entidades o materias, evitando ambigüedades derivadas de fenómenos lingüísticos como la sinonimia y la homonimia, así como de las distintas formas en que un mismo nombre puede aparecer en los registros (Garzón-Farinós & Peset, 2018).

El presente trabajo se inserta en este diálogo sobre el uso de datos enlazados para documentar y compartir información del patrimonio cultural, un ámbito en el que ha crecido la preocupación por la consistencia y calidad de los datos. Se propone como objetivo presentar un modelo de metadatos para la descripción de autoridades de personas en Wikidata, que incluye una plantilla de propiedades estandarizadas con sus valores, sugerencias de registro y niveles de obligatoriedad de uso. Este modelo busca mejorar la calidad, consistencia y reutilización de la información, facilitando su aplicación en contextos de gestión bibliográfica, archivística y de datos abiertos, y contribuyendo al fortalecimiento de prácticas de modelado de datos en entornos colaborativos.

Metodología

Para la creación de la plantilla de propiedades en Wikidata, se siguió un proceso sistemático orientado a estandarizar la descripción de datos de autoridades de personas. Esto incluyó las siguientes etapas:

- Búsqueda y examen crítico de la literatura existente en torno al control de autoridad y los catálogos de autoridad.
- Revisión de mejores prácticas en el modelado de datos para personas en Wikidata:

Se escogieron dos proyectos destacados liderados por organizaciones reconocidas internacionalmente, para analizar el uso de propiedades para la descripción de personas. El

primero de ellos ha sido *Smithsonian Research Online*, liderado por las bibliotecas de la *Smithsonian Institution*, que explora cómo los investigadores de dicha institución están representados en Wikidata. En su página de proyecto en Wikidata (Wikidata:WikiProject PCC Wikidata Pilot/Smithsonian Libraries/Projects/Smithsonian Research Online) se incluyen las propiedades básicas y las propiedades extendidas para la descripción de los investigadores afiliados a la institución. Entre las propiedades básicas se destacan "instancia de" (P31); "campo de trabajo" (P101), "empleador" (P108) e identificadores como el ORCID ID y el ISNI. Entre las propiedades extendidas, se destacan "ocupación" (P106); "fecha de nacimiento" (P569); "lugar de nacimiento" (P19); "fecha de fallecimiento" (P570); "lugar de fallecimiento" (P20); "educado en" (P69) junto con varios cualificadores; "premio recibido" (P166) y "obra destacada" (P800).

El segundo de los proyectos revisados fue el Wikidata:WikiProject Visibilizando el dominio público colombiano, que se enfoca en la creación de un directorio sobre autores colombianos cuyas obras son parte del dominio público. Entre las propiedades destacadas se encuentran "sexo o género" (P21); "país de nacionalidad" (P27); "nombre de pila" (P735); "apellido" (P734); "segundo apellido" (P1950); "seudónimo" (P742) y "estado de los derechos de autor como creador" (P7763).

- Evaluación de las políticas para la creación de elementos de personas en Wikidata:

En esta etapa se revisó la política de Wikidata sobre la creación de datos para personas vivas (Wikidata: Living people), que aplica a todas las personas nacidas durante los últimos 115 años. Esta política establece que el respeto a la dignidad y privacidad de las personas debe guiar la creación de datos. En este sentido, los datos provistos no deben violar las expectativas básicas de privacidad y deben sustentarse en información pública o abiertamente ofrecida por el individuo en cuestión. Para guiar la implementación de la esta política, Wikidata ha compilado listas con propiedades que pueden resultar controversiales y propiedades que podrían atentar contra la privacidad de las personas. Destacamos que varias propiedades utilizadas comúnmente en la descripción de personas, tales como ocupación, empleador o lugar de nacimiento, están incluidas en esta lista. Esto no implica que no se puedan utilizar, pero sí hay que validar que estos datos sean tomados de fuentes abiertas al público. Además, el uso de estos datos debe complementarse con la inclusión de referencias que apunten a la fuente de la cual se obtuvo el dato. Las redes sociales personales no son fuentes confiables ni necesariamente se deben considerar como fuentes públicas.

- Análisis comparativo de las mejores prácticas y políticas para la creación de autoridades de nombres según la norma catalográfica *Resource Description and Access* (RDA):

Como parte de la cuarta etapa, se revisaron los elementos para la descripción de personas según la norma RDA, el estándar internacional que guía la descripción bibliográfica y la creación de registros de autoridades para personas en el contexto de las bibliotecas. Según Hernández Sánchez (2022) RDA cuenta con muchos elementos que se pueden emplear al construir catálogos de autoridad de nombres personales. A continuación, se mencionan los elementos junto con sus etiquetas MARC: forma autorizada del nombre (100); variantes del nombre (400); fecha de nacimiento o fallecimiento (046, subcampos $f y $g respectivamente); lugares relacionados con la persona, tales como lugar de nacimiento, lugar de fallecimiento, país relacionado, lugar de residencia (370, subcampos $a, $b, $c y $e, respectivamente); disciplinas o áreas de especialidad (372); afiliaciones institucionales (373); ocupación (374); género (375); idiomas relacionados (377) y otros atributos de la persona (368).

- Armonización de los lineamientos identificados:

Como último procedimiento metodológico para la selección y estandarización de las propiedades, según se muestra en la Tabla 1, también se llevó a cabo un mapeo entre los elementos de RDA y las propiedades en Wikidata para la descripción de personas, para identificar propiedades comunes.

ELEMENTOS RDA (MARC)	PROPIEDADES EN WIKIDATA
punto de acceso autorizado para la persona (100 $a)	etiqueta nombre de pila (P735) apellido (P734) segundo apellido (P1950)
variante de nombre de persona (400 $a)	también conocido como Nombre corto (P1813)
identidad alterna de persona (500 $a)	seudónimo (P742)
fecha de nacimiento (046 $f) fecha de fallecimiento (046 $g)	fecha de nacimiento (P569) fecha de fallecimiento (P20)
lugar de nacimiento (370 $a) lugar de fallecimiento (370 $b) lugar de residencia (370 $e) país asociado con persona (370 $c)	lugar de nacimiento (P19) lugar de fallecimiento residencia (P551) no hay equivalencia exacta. Se establece como "país de nacionalidad" (P27)
campo de actividad de persona (372 $a)	campo de trabajo (P101)
organismo corporativo relacionado de persona (373 $a) empleador (373 $a / 510 $a) graduado de (373 $a / 510 $a)	afiliación (P1416) empleador (P108) educado en (P69)
profesión u ocupación (374 $a)	ocupación (P106)
género (375 $a)	sexo o género (P21)

ELEMENTOS RDA (MARC)	PROPIEDADES EN WIKIDATA
lenguaje de persona (377 $a)	lenguas habladas, escritas o signadas (P1412)
identificador para persona (010, 024)	identificador VIAF (P214) identificador ISNI (P213) identificador de autoridades de la Biblioteca del Congreso de EEUU (P244) ORCID iD (P496) ResearcherID (P1053)

Tabla 1. Correspondencia entre elementos RDA (codificados en formato MARC) y propiedades de Wikidata para la descripción de personas.

Importante resaltar que esta lista presentada en la Tabla 1 no es exhaustiva. Se trata de una síntesis, y su propósito es la identificación de propiedades básicas que deben ser consideradas al momento de describir personas.

Las cinco etapas metodológicas descritas permitieron identificar, comparar y armonizar un conjunto de propiedades relevantes para la descripción de autoridades de personas en Wikidata, integrando buenas prácticas provenientes de la literatura, proyectos aplicados, políticas de la plataforma y estándares bibliográficos internacionales. A partir de este proceso de sistematización y análisis comparativo, se construyó una base para la definición de un modelo de metadatos estandarizado, el cual se presenta en la siguiente sección como resultado principal de este estudio.

Propuesta de modelado

La Tabla 2 muestra la propuesta para el modelado de los metadatos de autoridades de persona en Wikidata. En total, se incluyen 23 propiedades, cada una con el valor que puede asignarse, una breve guía para su descripción y el nivel de obligatoriedad de su uso (requerido, recomendable u opcional). Estos campos permiten describir de manera estructurada los aspectos biográficos, académicos y bibliográficos de cada individuo, incluyendo su identidad, formación, ocupacional, afiliaciones, obras destacadas y enlaces a identificadores externos.

PROPIEDAD	VALOR	GUÍA PARA LA DESCRIPCION	USO
etiqueta		Nombre de la persona	Requerido
descripción		Breve descripción en letras minúsculas	Requerido
variantes del nombre		Otras formas del nombre	Requerido
instancia de (P31)	ser humano (Q5)		Requerido

PROPIEDAD	VALOR	GUÍA PARA LA DESCRIPCION	USO
nombre de pila (P735)	nombre	Se puede repetir para registrar nombres compuestos.	Requerido
apellido (P734)	apellido paterno		Requerido
segundo apellido (P1950)	apellido materno		Opcional
sexo o género (P21)*	masculino (Q6581097), femenino (Q6581072), intersexual (Q1097630), mujer transgénero (Q1052281), hombre transgénero (Q2449503), no binario (Q48270), agénero (Q505371), genderqueer (Q12964198), género fluido (Q18116794), neutro (Q1289754)	Seleccionar según información confiable	Opcional*
fecha de nacimiento (P569)		Incluir si la información está disponible en fuentes confiables.	Opcional
lugar de nacimiento (P19)		Incluir si la información está disponible en fuentes confiables.	Opcional
país de nacionalidad (P27)*			Opcional*
profesión u ocupación (P106)	académico (Q3400985), profesor universitario (Q1622272), investigador (Q1650915), otros según especialidad académica	Se puede repetir para registrar distintos valores.	Requerido
empleador (P108)	Institución universitaria	Se puede repetir para registrar distintos valores. Puede incluir cualificadores tales como fecha de comienzo y terminación.	Requerido
campo de trabajo (P101)	Disciplina académica	Se puede repetir para registrar distintos valores.	Requerido
afiliación (P1416)	Departamento académico	Se puede repetir para registrar distintos valores. Puede incluir cualificadores tales como fecha de comienzo y terminación.	Opcional

PROPIEDAD	VALOR	GUÍA PARA LA DESCRIPCION	USO
educado en (P69)	Institución universitaria	Se puede repetir para registrar distintos valores. Puede incluir cualificadores tales como fecha de comienzo y terminación y grado académico.	Opcional
obra destacada (P800)	Publicaciones, obras destacadas	Ej.: libros, artículos,	Opcional
lenguas habladas, escritas o signadas (P1412)	lenguas habladas por la persona o en la que escribe sus obras		Opcional
VIAF ID (P214)			Recomendable, si disponible
identificador de autoridades de la Biblioteca del Congreso de EE. UU. (P244)		Identificador para los nombres que forman parte del archivo de autoridades de la Biblioteca del Congreso.	Recomendable, si disponible
Identificador ISNI (P213)		*International Standard Name Identifier* es un identificador único para personas y organizaciones involucradas en la creación de contenido creativo (autores, músicos, investigadores, artistas, editoriales, etc.).	Recomendable, si disponible
identificador ORCID (P496)		Identificador para personas que tienen un perfil profesional en ORCID.	Recomendable, si disponible
ID Researcher (P1053)		Identificador de investigadores introducido en 2008 por Thomson Reuter	Recomendable, si disponible

Tabla 2. Propuesta modelo de metadatos de autoridades en Wikidata

* Las propiedades *sexo o género* (P21) y *país de nacionalidad* (P27) han sido ampliamente señaladas en la literatura como especialmente críticas dentro del modelo de Wikidata, por lo que su uso requiere una atención particular y una validación cuidadosa de las fuentes.

Discusión

El control de autoridades constituye un componente fundamental en la organización y normalización de los datos, ya que contribuye a la identificación unívoca de los recursos. Tradicionalmente, este proceso ha proporcionado consistencia al catálogo bibliográfico en contenido y forma (Garzón-Farinós & Peset, 2018). Sin embargo, los cambios tecnológicos influyen en el tratamiento de la información, así como en su presentación, almacenamiento y recuperación. En consecuencia, a medida que aumenta el volumen de la información en internet, el control de autoridades se hace cada vez más necesario. En el caso de los datos enlazados y las plataformas colaborativas como Wikidata, esta relevancia se intensifica aún más, debido a la naturaleza abierta y distribuida de la creación de los metadatos.

En este escenario, se han identificado inconsistencias estructurales en Wikidata, particularmente en el uso de propiedades como *instancia de* y *subclase de*. Brasileiro et al. (2016) evidencian que el uso inadecuado de estas propiedades genera patrones de error recurrentes que afectan la coherencia de los modelos de clasificación y la representación del conocimiento. Estos hallazgos se vinculan con una problemática más amplia relacionada con la ausencia de criterios uniformes en la descripción de los datos dentro de la plataforma.

Asimismo, considerando que no existe una única forma de modelar la información (Pastor-Sánchez, 2021), los desafíos identificados en Wikidata se relacionan directamente con su naturaleza abierta y colaborativa. Trucco y Blanco (2024) señalan que la multiplicidad y ambigüedad de algunas propiedades evidencian la falta de especificaciones claras sobre su uso. Aunque la Fundación Wikimedia establece lineamientos generales y la comunidad participa activamente en la toma de decisiones, la ausencia de reglas estrictamente normalizadas genera heterogeneidad en la práctica de edición, lo que impacta la calidad y consistencia de los datos. En esta línea, Vrandečić y Krötzsch (2014) reconocen desafíos relacionados con la consistencia, la estandarización y la gestión de información contradictoria.

No obstante, estas limitaciones coexisten con ventajas significativas de la plataforma. Wikidata ofrece capacidades que superan a los sistemas tradicionales de control de autoridades, como la integración de múltiples fuentes de información, la interoperabilidad multilingüe y su contribución a la construcción de un grafo de conocimiento global (Bianchini & Sardo, 2022). En este sentido, Agenjo-Bullón y Hernández-Carrascal (2018) han señalado que Wikidata se perfila como un referente emergente en la gestión de autoridades, especialmente por su capacidad de conectar datos de forma abierta y dinámic.

A partir de ello, el modelo propuesto en este estudio busca reducir dicha heterogeneidad mediante la definición de un conjunto mínimo de propiedades y criterios de uso para la normalización de elementos de nombres personales. Este enfoque se concreta en 23 propiedades alineadas con prácticas empleadas en bibliotecas académicas internacionales. Sandberg y Jin (2016) destacan que la calidad del control de autoridades mejora cuando se vinculan múltiples identificadores externos, como VIAF, ORCID, ISNI y Scopus, lo que permite una mayor fiabilidad que el uso de una única fuente. Esta aproximación refuerza la pertinencia de modelos híbridos de autoridad en entornos como Wikidata, donde la interoperabilidad y la diversidad de fuentes constituyen elementos centrales.

Tal como muestran Thornton et al. (2022), el proceso de mapear datos de investigación a propiedades de Wikidata implica decisiones de modelado que afectan directamente la representación del conocimiento. La elección de propiedades como P21 (sexo o género), P106 (profesión u ocupación) o P19 (lugar de nacimiento) no solo estructura la información, sino que también determina cómo los sujetos son descritos y conectados dentro del grafo global.

La propiedad P21 (sexo o género) en Wikidata ha sido ampliamente discutida en la literatura académica, ya que no representa un único concepto, sino que mezcla sexo asignado al nacer, identidad de género, expresión de género y modalidad de género. Estudios como el proyecto WiGeDi han señalado que esta ambigüedad puede generar problemas de representación, especialmente en el caso de personas trans y no binarias. Asimismo, la propiedad P27 (país de nacionalidad) dificulta la expresión de la identidad cultural y no toma en cuenta la complejidad de la geopolítica colonial (Baker y Mahal, 2024). Además, está propiedad no es adecuada para registrar los usos premodernos de la idea de ciudadanía o el surgimiento histórico de estados soberanos, lo cual puede resultar en representaciones anacrónicas.

Conclusiones

El presente trabajo ha puesto de manifiesto que, si bien Wikidata ofrece un entorno flexible y en constante evolución para la representación de datos, la ausencia de esquemas de modelado uniformes y estrictamente definidos genera inconsistencias en la descripción de autoridades de personas, afectando la calidad, recuperación e interoperabilidad de la información.

Una vía para lograr descripciones más estandarizadas es la elaboración de mapeos con otros estándares de metadatos. El *Resource Description and Access* (RDA), implementado en el ámbito bibliotecológico, ofrece un modelo normalizado para la descripción de personas, con elementos que describen las características de una persona, así como sus relaciones con otras

entidades como fechas, lugares e instituciones. A partir de ello, es posible establecer relaciones de equivalencia entre elementos RDA y propiedades en Wikidata, de manera que se facilite la consistencia en la descripción, sin renunciar a la incorporación de nuevas propiedades.

Como principal aporte, este trabajo propone un modelo de metadatos que integra propiedades de Wikidata y está acompañado de una clasificación por niveles de uso (requerido, recomendable y opcional). Este modelo contribuye a: (1) mejorar la consistencia en la descripción de personas, (2) facilitar la integración con sistemas bibliográficos tradicionales y (3) promover la reutilización de datos en entornos de datos abiertos enlazados. Asimismo, la correspondencia establecida entre RDA y Wikidata constituye una base metodológica útil para futuros procesos de armonización entre estándares.

Si bien es cierto que la descripción o modelado de datos es un proceso técnico, no se debe perder de vista la dimensión social inherente a dicha tarea. Por lo tanto, resulta fundamental tomar en cuenta asuntos como la privacidad de las personas y las ideologías implícitas en las ontologías, ya que esto tiene un impacto tanto en el acceso y recuperación de la información como en la interpretación y creación de conocimiento. Como líneas futuras de investigación, se propone validar este modelo mediante su aplicación en conjuntos de datos reales, así como explorar su integración en proyectos del ámbito Open GLAM que utilicen Wikidata como infraestructura de datos.

Referencias

Agenjo-Bullón, X. y Hernández-Carrascal, F. (2018). Registros de autoridades, enriquecimiento semántico y Wikidata. *Anuario ThinkEPI*, *12*, 361–372. https://doi.org/10.3145/thinkepi.2018.61

Arroyo-Machado, W. (2024). Explorando Wikidata para la investigación en Ciencias Sociales. *Infonomy, 2*(2), e24034. https://doi.org/10.3145/infonomy.24.034

Baker, J. y Mahal, A. K. (2024). I have always found the whole area a minefield: Wikidata, historical lives, and knowledge infrastructure. *International Journal of Digital Humanities, 6*, 217–236. https://doi.org/10.1007/s42803-024-00090-5

Bianchini, C. y Sardo, L. (2022). Wikidata: A new perspective towards universal bibliographic control. *JLIS.it, 13*(1), 291–311. https://doi.org/10.4403/jlis.it-12725

Brasileiro, F. Almeida, J. P. A. Carvalho, V.A. y Guizzardi, G. (2016). Applying a Multi-Level Modeling Theory to Assess Taxonomic Hierarchies in Wikidata. En J. Bourdeau, J. A. Hendler y

R. Nkambou Nkambou (Coords.), *Proceedings of the 25th International Conference Companion on World Wide Web* (pp. 975–980). International World Wide Web Conferences Steering Committee. https://doi.org/10.1145/2872518.2891117

Farda-Sarbas, M. y Müller-Birn, C. (2019). Wikidata from a Research Perspective - A Systematic Mapping Study of Wikidata. *arXiv,* 1908.11153. https://doi.org/10.48550/arXiv.1908.11153

Garzón-Farinós, F. y Peset, F. (2018). Normativas interoperables sobre control de autoridades para bibliotecas digitales semánticas. *Investigación Bibliotecológica*, *32*(76), 167-188 https://doi.org/10.22201/iibi.24488321xe.2018.76.58018

Herrera Cubides, J. F. Gaona García, P. y Sánchez Alonso, S. (2018). Linked Data: qué sucede con la heterogeneidad y la interoperabilidad. *Scientia et Technica, 23*(2), 230–240. https://doi.org/10.22517/23447214.16791

Hernández Sánchez, N. (2022). Los catálogos de autoridad de nombres personales: propósitos y metadatos. *Investigación Bibliotecológica*, *36*(92), 193–211. https://doi.org/10.22201/iibi.24488321xe.2022.92.58587

Pastor-Sánchez, J.A. (2021). Wikidata como herramienta para elaborar ontologías y vocabularios controlados. *Anuario ThinkEPI*, *15, e15f0.* https://doi.org/10.3145/thinkepi.2021.e15f01

Piscopo, A. Phethean, C. y Simperl, E. (2017). What makes a good collaborative knowledge graph: Group composition and quality in Wikidata. En E. Ferrara, O. Varol, F. Menczer y A. Flammini (Eds.), *Social Informatics (SocInfo 2017), Part I* (pp. 305–322). Springer. https://doi.org/10.1007/978-3-319-67217-5_19

Ripoll, M. Reeves, N. Kurteva, A. Simperl, E. Meroño Peñuela, A. y Diepold, K. (2025). Filling in the blanks? A systematic review and theoretical conceptualisation for measuring Wikidata content gaps. arXiv, 2505.16383v2 https://doi.org/10.48550/arXiv.2505.16383

Sandberg, J., y Jin, Q. (2016). How should catalogers provide authority control for journal article authors? Name identifiers in the linked data world. *Cataloging & Classification Quarterly, 54*(8), 537–552. https://doi.org/10.1080/01639374.2016.1238429

Sorensen, A. H. Menyuk, R. Shapiro, D. y Naples, R. (2026). Piloting Wikidata as an authority identifier: The (In)visible Women Project at the Smithsonian Institution. *Journal of Open Humanities Data, 12*(22), 1–8. https://doi.org/10.5334/johd.467

Thornton, K. Seals-Nutt, K. Van Remoortel, M. Birkholz, J. M. y De Potter, P. (2022). Linking women editors of periodicals to the Wikidata knowledge graph. *Semantic Web: Interoperability, Usability, Applicability, 14*(2), 443-455. https://doi.org/10.3233/SW-222845

Tharani, K. (2021). Much more than a mere technology: A systematic review of Wikidata in libraries. *The Journal of Academic Librarianship, 47*(2), 102326. https://doi.org/10.1016/j.acalib.2021.102326

Trucco, C. J. y Blanco, M. E. (2024). *Descripción y acceso abierto: Articulación entre ISAD-G y metadatos Wikidata para fotografías digitalizadas*. Redes. https://editorialredes.omeka.net/items/show/31

Vrandečić, D. y Krötzsch, M. (2014). Wikidata: A free collaborative knowledgebase. *Communications of the ACM, 57*(10), 78–85. https://doi.org/10.1145/2629489

Vellucci, S. L. (2000). Metadata and authority control. *Library Resources & Technical Services, 44*(1), 33–42. https://doi.org/10.5860/lrts.44n1.33

Waagmeester, A. Stupp, G. Burgstaller-Muehlbacher, S. Good, B. M. Griffith, M. Griffith, O. L., Hanspers, K. Hermjakob, H. Hudson, T. S. Hybiske, K. Keating, S. M. Manske, M. Mayers, M. Mietchen, D. Mitraka, E. Pico, A. R. Putman, T. Riutta, A. Queralt-Rosinach, N. … Su A. I. (2020). Wikidata as a knowledge graph for the life sciences. *eLife, 9*, e52614. https://doi.org/10.7554/eLife.52614

Zhu, L. (2019). The future of authority control: Issues and trends in the linked data environment. *Journal of Library Metadata, 19*(3–4), 215–238. https://doi.org/10.1080/19386389.2019.1688368