

Metodología de trabajo colaborativo en Wikidata

Creación de perfiles de personas autoras colombianas en el dominio público para Colombia

Fundación Conector David Ramírez-Ordóñez
Leonardo RamírezOrdóñez Martha Muñoz

2024-02-15

Metodología de trabajo colaborativo en Wikidata:

Creación de perfiles de personas autoras colombianas en el dominio público para Colombia

Resumen

Como parte del proyecto ‘Visibilizando el dominio público colombiano’, a continuación se presenta la explicación de los pasos necesarios para la creación de elementos en Wikidata de autores de Colombia en dominio público. Se trata de la formulación de una estrategia de búsqueda y recopilación de fuentes de información, su organización en una hoja de cálculo, la validación de esta información y la publicación en la base de datos semántica Wikidata. ¹.

Introducción

Este documento explica los pasos necesarios para la creación de elementos en Wikidata de autores de Colombia en dominio público, basado en estrategias de búsqueda y recopilación de fuentes de información, su organización en una hoja de cálculo, la validación de esta información y la publicación en Wikidata. El documento segmenta todo este proceso en tres partes: primero, la metodología de la creación de la estrategia de búsqueda y la recopilación de información. La segunda parte, sobre la publicación en Wikidata de los datos recopilados mediante una guía. La tercera parte contiene ideas para difundir el trabajo realizado y consideraciones finales.

El dominio público en Colombia se empezó a celebrar desde el 1 de enero de 2016, mediante una publicación en el sitio web de la Biblioteca Nacional de Colombia (Biblioteca Nacional de Colombia, 2016). Este listado es la fuente oficial de los datos de las personas autoras que pasan en dominio público en Colombia, a partir de las obras que se encuentran en el catálogo de la Biblioteca. Uno de

¹Este documento está en dominio público desde su creación. Puede contactarnos en nuestro sitio web: conector.red/

los principales retos es que dicha publicación no se encontraba en un formato que cumpliera el estándar de las 5 estrellas de los datos abiertos enlazados (Hausenblas, 2012). Por lo tanto, el objetivo final es la publicación de los datos de personas autoras cumpliendo estándares de datos abiertos enlazados mediante Wikidata.

Por otra parte, entendemos que la identificación de personas en dominio público debería ser un trabajo distribuido. Es decir, si bien la iniciativa de la Biblioteca Nacional de Colombia es muy valiosa y necesaria, mucho mejor si hay más personas e instituciones colaborando en la visualización del dominio público colombiano.

Primera parte: metodología

Definir la estrategia de búsqueda de información

Nuestra estrategia de búsqueda tiene tres partes: la identificación de variables, datos y los pasos a seguir para generar un listado de personas autoras en dominio público.

Las variables La estrategia de búsqueda deberá identificar las siguientes variables:

- La legislación de derecho de autor actualizada sobre cuándo una obra pasa a dominio público.
- Los diagramas de flujo del paso de una obra a dominio público. Si existe una calculadora de dominio público que lo ejemplifique, ésta puede ser de ayuda (Ramírez-Ordóñez, 2011).
- Las oficinas de registro de obras. Para el caso colombiano, contamos con el depósito legal de la Biblioteca Nacional de Colombia, la Biblioteca del Congreso de Colombia y la Biblioteca de la Universidad Nacional de Colombia (Biblioteca Nacional de Colombia, n.d.-b). También la oficina de registro de la Dirección Nacional de Derecho de Autor (Ministerio del Interior y de Justicia, República de Colombia, 2002) y las oficinas de registros de identificadores únicos, como ISBN, ISSN, ISMN y similares. Opcionalmente, la oficina de registro de identificación, para el caso colombiano, la Registraduría Nacional del Estado Civil.
- Las bibliotecas patrimoniales en donde puedan existir ejemplares de las obras.
- Los índices nacionales o sistemas de autoridades de autor (Biblioteca Nacional de Colombia, n.d.-a). Éstos deben contar con la validación de órganos oficiales.
- Los portales de datos abiertos nacionales, en donde puedan existir datos sobre las obras que pasan a dominio público anualmente.
- Las publicaciones oficiales donde se anuncien las obras que pasan a dominio público en el país. En el caso colombiano, la celebración del día del dominio público de la Biblioteca Nacional de Colombia.
- Las organizaciones o instituciones interesadas en la promoción de contenido libre, como los capítulos nacionales de Wikimedia. En nuestro caso, Wikimedia Colombia (*Wikimedia Colombia*, 2023).

Los datos Los datos que buscamos son los relacionados a las personas autoras, principalmente y secundariamente las de sus obras. Por esto buscamos los siguientes datos:

- Nombres y apellidos.
- Fecha de nacimiento.
- Fecha de fallecimiento.
- Fecha en la que su obra pasa a dominio público, de acuerdo a la legislación nacional.
- Lugar de nacimiento (opcional).
- Nacionalidad.
- Pseudónimos (opcional).
- Género.
- Lugar de fallecimiento.
- Obras destacadas. Que son definidas por las instituciones patrimoniales.
- Institución en donde puede consultarse su obra.
- Identificadores únicos (VIAF, ISNI, etc.) (opcional).

Sobre las obras, los datos que buscamos son:

- Título.
- Año de publicación.
- Persona que creó la obra.
- Trabajo completo disponible en URL. (Opcional)

Los pasos Con las variables y los datos a solicitar definidos, procedemos a definir la estrategia de búsqueda de información. En los anexos a este documento podrás encontrar los formatos con los campos de recopilación de información.

- Identifica el periodo de tiempo que quieres cubrir.
- Identifica las fuentes de información que produzcan listados oficiales de personas autoras que pasan a dominio público.
- Identifica las instituciones en donde se pueden consultar dichas obras.
- Diligencia los formatos con los datos del autor y sus obras destacadas.

Realizar la búsqueda de información

Una vez encuentres los listados de autores en dominio público, puedes comprobar sus datos en el catálogo de la biblioteca que genera los listados. Idealmente, los catálogos tendrán información de las fechas de nacimiento y muerte de las personas autoras.

Los datos de las personas autoras pueden corroborarse con otras instituciones patrimoniales y sus catálogos. En nuestro caso, hemos encontrado estas fuentes de información como útiles:

- El listado de obras en dominio público de la Biblioteca Nacional de Colombia. Aquí el ejemplo de 2022.
- La enciclopedia de Banrepcultural
- El catálogo de la Biblioteca Nacional de Colombia
- Los catálogos de la red de bibliotecas del Banco de la República.
- Catálogo de la Biblioteca Pública Piloto de Medellín.

- La biblioteca de la Academia Colombiana de Historia. Los departamentos y sus casas de la cultura también tienen información. Por ejemplo:
 - El repertorio histórico de la Academia Antioqueña de Historia
- Repositorios de universidades como el de
 - la Universidad Santo Tomás de Aquino
 - La biblioteca de la Universidad de Antioquia
- La Radio Nacional de Colombia
- Señal Memoria
- VIAF

Diseñar el perfil de aplicación

Lo que llamamos un perfil de aplicación es un esquema de datos que se usarán en la identificación de las personas en dominio público y sus obras. Piense en qué datos debería conocer sobre una persona para identificarla y la respuesta es lo que debe contener el esquema. Datos como su nombre, su lugar y fecha de nacimiento o el título de la obra que realizó.

En realidad se requieren dos perfiles para esta aplicación: la de las personas y la de sus obras. Presentaremos los diccionarios de datos para cada uno de estos dos perfiles de aplicación. Creamos esta hoja de cálculo con el esquema de las personas autoras y otra hoja de cálculo con el esquema de las obras.

Usaremos diccionarios de datos que tengan en cuenta los datos de Wikidata, para poder identificar el resultado final como guía. Por eso, usamos a Porfirio Barba Jacob (*Porfirio Barba-Jacob*, n.d.) y su obra “La canción de la vida profunda y otros poemas” (Barba-Jacob, n.d.) como ejemplos.

Diccionario de datos para personas El diccionario de datos de personas autoras se presenta en la siguiente tabla:

Nombre del campo	Valor
ID-Interno	Proviene de la pestaña “Persona natural sin duplicados (199)”. Es el número consecutivo que identifica a un autor personal en nuestro proyecto. Ejemplo: el ID-Interno 148
Qid	corresponde a Barba Jacob, Porfirio Identificador de Wikidata de una persona. Ejemplo: el Qid Q1744460 corresponde a Porfirio Barba Jacob en Wikidata.
P31 (instancia de)	El tipo de objeto del elemento en Wikidata. En nuestro caso, siempre será Ser humano
Instancia de (Q)	Identificador de Wikidata de Ser humano. Siempre es Q5

Nombre del campo	Valor
P18 (imagen)	URL de la imagen en Wikimedia Commons de la persona descrita. Ejemplo: https://commons.wikimedia.org/wiki/File:Porfirio_Barba_Jacob-Busto-Medellin.JPG
Imagen (Nombre del archivo en Commons)	Nombre del archivo de la imagen en Commons. Ejemplo: File:Porfirio_Barba_Jacob-Busto-Medellin.JPG
sexo o género (P21)	Nombre del sexo de la persona. Los posibles valores son: masculino, femenino, intersexual, mujer transgénero, hombre transgénero, no binario; Ejemplo: Masculino
Sexo o género (Q)	Para humanos: masculino (Q6581097), femenino (Q6581072), intersexual (Q1097630), mujer transgénero (Q1052281), hombre transgénero (Q2449503), no binario (Q48270); Ejemplo de “masculino”: Q6581097
país de nacionalidad	Estado soberano del cual la persona es ciudadana. En nuestro caso, generalmente es Colombia
País de nacionalidad (Q)	Valor en Wikidata del país de nacionalidad. Ejemplo para describir a “Colombia”: Q739
nombre de pila (P735)	Todos los nombres de la persona. Ejemplo: Miguel Ángel (Q40440281). No se deben registrar los nombres por separado.
apellido	Todos los apellidos de la persona. Ejemplo: Osorio Benítez
Apellido 1 (Q)	Identificador de Wikidata del primer apellido. En nuestro caso, el ejemplo de Osorio: Q37483500
Apellido 2 (Q)	Identificador de Wikidata del apellido materno en español (P1950). En nuestro caso, el ejemplo de Benitez: Q26736553
seudónimo	alias usado por alguien o por el cual esta persona es conocida, como texto libre. En nuestro ejemplo: Porfirio Barba Jacob

Nombre del campo	Valor
nombre en el idioma nativo	nombre de una persona en su idioma nativo. En nuestro ejemplo: Porfirio Barba Jacob (Latin American Spanish)
fecha de nacimiento (AAAA-MM-DD)	fecha en la cual nació la persona.
lugar de nacimiento	sitio más específico conocido donde nació la persona
fecha de fallecimiento	fecha de defunción
lugar de fallecimiento	sitio donde murió la persona
cónyuge	el sujeto tuvo como cónyuge a
Cónyuge (Q)	La pareja debe crearse en Wikidata para poder incluirla en este campo.
ocupación	actividad laboral u otra ocupación de la persona. Ya que una persona puede tener múltiples campos, este valor puede repetirse cuantas veces sea necesario. En nuestro ejemplo: Poeta, escritor
Ocupación 1 (Q)	actividad laboral u otra ocupación de la persona. Una de las repeticiones.
Ocupación 2 (Q)	actividad laboral u otra ocupación de la persona. Una de las repeticiones.
lenguas habladas, escritas o signadas	idioma que habla, escribe o signa una persona, incluyendo el idioma materno o nativo
lengua de escritura	lengua en la cual el escritor ha escrito su obra
obra destacada	obra científica, literaria o de arte destacada del sujeto. Debe estar creada en Wikidata para poder incluirla en este campo.
tiene obras en la colección	colección que tiene obras de esta persona. Este valor puede repetirse cuantas veces sea necesario.
Tiene obras en la colección 1 (Q)	colección que tiene obras de esta persona. Una de las repeticiones.
Tiene obras en la colección 2 (Q)	colección que tiene obras de esta persona. Una de las repeticiones.
estado de los derechos de autor como creador	indica si las obras de este creador siguen protegidas por las leyes de derechos de autor, o ya están en el dominio público. En nuestro ejemplo: copyrights on works have expired.
página web oficial	URL de la página de web oficial de este elemento
Educado en (P69)	instituciones educativas en las que estudió.

Nombre del campo	Valor
categoría en Commons (Nombre de la categoría)	nombre de la categoría de Wikimedia Commons con archivos relacionados con este elemento (sin el prefijo «Categoría»)
identificador ORCID	identificador ORCID de una persona
identificador de autoridades de la Biblioteca del Congreso de EE. UU. (Id)	número identificativo proporcionado por la Biblioteca del Congreso de los Estados Unidos
identificador VIAF (Id)	Virtual International Authority File identifier
ISNI (Id)	International Standard Name Identifier
Observaciones	Notas internas para registrar información sobre la persona. No se publicarán en Wikidata, pero ayudan a compilar recursos, fuentes bibliográficas o dudas que se puedan tener en el proceso de creación de este elemento en Wikidata.

Diccionario de datos para obras El diccionario de datos de las obras se presenta en la siguiente tabla:

Nombre del campo	Valor
instancia de	El tipo de objeto del elemento en Wikidata. En nuestro caso, puede ser obra escrita, composición musical u otras.
autor	Quien crea la obra. Para obras musicales, use P86 - Compositor.
nombre del autor	cadena (string) para guardar el nombre del autor cuando éste no se encuentra en Wikidata. En nuestro caso no aplica, porque siempre usaremos autores de Wikidata.
editor	editor de una obra compilada o de una publicación periódica como un periódico o una publicación académica
título	título original
subtítulo	subtítulo de la obra que sigue al título
idioma de la obra o del nombre	idioma de la obra, edición o traducción
serie	serie de obras
precedido por	inmediato predecesor en alguna serie de la que el elemento forma parte

Nombre del campo	Valor
sucedido por	Inmediato sucesor en alguna serie de la cual forma parte el elemento.
editorial	organización responsable de la publicación de textos o software. Wikidata pide que creamos este elemento si no existe.
fecha de publicación	fecha en la cual fue publicada esta obra
lugar de publicación	ubicación geográfica de la edición (usa el lugar de la 1ª edición para las obras)
tema principal de la obra	materia de la que trata
género	género literario
lugar de la narración	la narración de esta obra se desarrolla en esta ubicación
edición o traducción	enlace a una edición de este elemento
país de origen	país de origen de la obra o el objeto
fecha de dominio público	fecha en la cual un elemento pasa a estar en el dominio público en una jurisdicción
forma artística	forma literaria
composer	persona encargada de escribir la música de una obra musical. Se usa en vez de “autor”.
trabajo completo disponible en URL	URL de la página web que contiene el texto completo de este elemento

Recopilar los datos

Los datos se recopilan en un archivo como la hoja de cálculo para subir datos. Este diccionario de datos se construyó pensando en dos columnas, la de lectura humana y la de lectura de máquina. En la lectura humana, se completa con un texto comprensible por humanos. En la lectura de máquina, el título va acompañado por una “Q” y presenta el valor en Wikidata que se representa, para familiarizarnos con el funcionamiento de Wikidata. Por otro lado, si el valor de un campo no existe, debe dejarse en blanco. Tenga en cuenta que cada dato que recopile debería tener una referencia bibliográfica que lo respalde.

Puedes empezar recopilando los nombres y apellidos de las personas autoras. Si tienes sus nombres completos, éstos pueden compararse con los datos existentes en Wikidata para evitar crear duplicidad. Ya que Wikidata se basa en referencias, mejor si los datos que incluyes tienen una referencia que sustente lo que declaras en cada dato.

El resultado de este proceso debería ser mínimo un listado con los nombres y apellidos de las personas en dominio público. También, sus fechas de nacimiento y muerte. Otro listado similar debería tener las obras de estas personas con título, año de publicación e información sobre la autoría. A partir de estos datos, lo que haremos será comparar si los datos existen o no en Wikidata. Si

existen, incluiremos la información que haga falta. Si no existen, crearemos los elementos en Wikidata que hagan falta. Esto es lo que contaremos en la segunda parte.

Segunda parte: La guía para cargar los datos

Las diferentes formas de comparar los datos

Hay diferentes maneras de subir los datos a Wikidata. Puedes subir los datos de forma manual o en lote. Cada una de estas formas tiene sus ventajas en términos de tiempo o precisión de la información.

Lo primero que debes hacer es crear una cuenta en Wikidata o acceder a ella si ya la tienes. Ahora bien, debes empezar obligatoriamente cargando datos manualmente para que subas de categoría. Las personas que ya tienen cierto número de colaboraciones en Wikidata, pueden cargar los datos en lote. Al momento de escribir esta guía, el mínimo de ediciones que se deben tener son 50. Debes convertirte en lo que en Wikidata se llama un *usuario auto confirmado*. En este sitio presentan los requisitos para ser usuario auto confirmado. Debes tener más de 50 ediciones y algunos días de antigüedad.

Puedes revisar si eres una persona auto confirmada en este sitio web. Debes diligenciar los campos de “Usuario” y al final de la página verás si lo eres en la sección “Cambios en los permisos”.

El proceso que realizaremos será verificar si los datos que tenemos de las personas en dominio público ya se encuentran en Wikidata o si debemos crear sus elementos. Para esto, se pueden realizar revisiones para hacer la carga de datos de forma manual o en lote. Para la carga de datos en lote se requieren herramientas externas que presentaremos a continuación.

Recuerda: la premisa es reutilizar los datos. La gracia de los datos abiertos enlazados no es que se dupliquen, sino que se enlacen. Si un elemento ya existe en Wikidata, la idea es completarlo con los datos que podamos aportar.

La carga de datos manual Para realizar una carga de datos manual, deberás acceder a cada uno de los elementos en Wikidata que representan a las personas en dominio público. Por ejemplo, aquí puedes consultar el elemento de Porfirio Barba-Jacob o el de Josefa Andrade Berti. La carga de información se hace mediante los botones de “Edición”. Existen tutoriales de Wikidata que pueden ser de ayuda.

Si realizas la carga de datos de forma manual, puedes probar Recoin. De esta manera, te aparecerá un indicador con la calidad de los datos del elemento que estas creando. También, las propiedades que otros valores similares usan para que completes el elemento que estás creando. La siguiente imagen muestra la calidad de los datos y las sugerencias de propiedades que podrían incluirse a una biografía. Una ventaja del uso de Recoin es que de acuerdo al elemento que se está editando, Recoin sugiere datos a completar.

La carga de datos en lote Ya que el proceso de carga de datos puede ser bastante repetitivo, se pueden usar herramientas que agilicen este proceso. Dos de

[español](#)
[Hiperterminal](#)
[Discusión](#)
[Preferencias](#)
[Beta](#)

Elemento [Discusión](#) Leer [Lista de seguimiento](#) [Contribuciones](#) [Salir](#)

[Más](#)

Josefa Andrade Berti (Q123975372)

Autora colombiana [editar](#)

[Josefa Andrade](#)

[Recoín: Propiedades más relevantes que faltan](#)

Property ID	Label	Relative	Add Claim
P19	lugar de nacimiento	61.96%	+
P69	educado en	23.43%	+
P373	categoría en Commons	16.03%	
P166	premio recibido	13.5%	+
P6886	lengua de escritura	11.53%	+
P1559	nombre en el idioma nativo	11.11%	
P108	empleador	6.97%	+
P103	lengua materna	6.23%	+
P463	miembro de	6.18%	+
P140	religión	5.82%	+

[En más idiomas](#)

[Configurar](#)

Idioma	Etiqueta	Descripción	También conocido como
español	Josefa Andrade Berti	Autora colombiana	Josefa Andrade

Figure 1: Ejemplo de uso de Recoín para carga manual de datos.

estas herramientas son QuickStatements (*QuickStatements*, n.d.) y OpenRefine (*OpenRefine*, 2010).

Actividad: completa los nombres, apellidos y género de las personas autoras que quieras visibilizar en Wikidata en la hoja de cálculo para subir datos en la columna “Nombre completo”. No olvides incluir su ID-Interno.

El uso de OpenRefine y QuickStatements para la carga de datos en lote

Revisaremos estas dos herramientas que te ayudarán a cargar la información de forma mucho más rápida a Wikidata. Primero, revisaremos QuickStatements para que sepas cómo funciona y tengas unas nociones básicas, pero luego nos centraremos en OpenRefine que es la herramienta que más usaremos.

QuickStatements QuickStatements sirve para cargar datos a Wikidata. QuickStatements es una herramienta que funciona en el navegador web. Es decir, no requiere instalarse. El reto está en entender la estructura de los datos que cargarás mediante esta herramienta. No te preocupes si puede ser muy técnico al inicio. Con la práctica, le puedes encontrar mucha utilidad.

El siguiente código actualiza datos mediante QuickStatements. Específicamente, agrega el campo “Tiene obras en la colección” del elemento María Luisa Aguado Martínez (Q16161225) con un enlace.

```
qid      has works in the collection Design Museum of Barcelona
reference URL  La url Q16161225 P6379 Q4889477 s854
"https://cataleg.museudeldisseny.cat/resultat/?search=&objectName=..."
```

Este tipo de códigos se ponen en el cajón de QuickStatements. Para ilustrar el proceso, veamos la creación de una persona en dominio público. El código usado es:

```
qid,P31,P21,P27,P735,P734,P734,P1412,P7763
,Q5,Q6581072,Q739,Q617272,Q15181057,Q21506799,Q1321,Q71887839
```

Y al cargarlo por QuickStatements, el resultado se ve así:



Lote en Wikidata por [Hiperterminal](#) [\[Lotes\]](#) [Discutir/revertir lote](#)
Estado: DONE

1 done Q123975372 [Q123975372]	CREATE Item	instance of [P31]:human [Q5] sex or gender [P21]:female [Q6581072] country of citizenship [P27]:Colombia [Q739] given name [P735]:Josefa [Q617272] family name [P734]:Andrade [Q15181057] family name [P734]:Berti [Q21506799] languages spoken, written or signed [P1412]:Spanish [Q1321] copyright status as a creator [P7763]:copyrights on works have expired [Q71887839]
--	-----------------------------	--

Figure 2: Creación de una persona en dominio público mediante QuickStatements

OpenRefine OpenRefine no sólo carga datos a Wikidata como lo hace QuickStatements, también sirve para limpiar los datos, compararlos y modificarlos.

Debes instalarlo en tu computador. Luego, puedes usar la hoja de cálculo que preparaste con la lista de personas en dominio público para hacer todas las ediciones en Wikidata desde OpenRefine.

Puedes descargar OpenRefine en este enlace. También, puedes ver tutoriales sobre cómo usar OpenRefine.

Conciliar los datos

- Prerrequisitos:
- instalar OpenRefine
- Tener cuenta en Wikidata
- Ser usuaria autoconfirmada

En la conciliación de datos lo que hacemos es verificar si los datos ya existen en Wikidata o deben crearse. Para esto, cargamos la hoja de cálculo con los nombres completos y el sexo de las personas a OpenRefine. Por supuesto, para usar este software debes instalarlo.

Cotejar datos En la pestaña de la columna “Nombre completo” selecciona **Cotejar > Iniciar cotejo**. Aparece una ventana que pide la variable a cotejar, es decir, le das pistas a Wikidata de qué tipo de elemento quieres recuperar. Selecciona **Humano Q5**, **Nombre de hombre Q12308941** o **Nombre de mujer Q11879590**. El resultado es que aparecerán algunos nombres cotejados si existen en Wikidata o para crear. La siguiente imagen lo ilustra. En azul aparecen los nombres que existen en Wikidata. En negro, los que no existen. Si aparecen varias opciones, deberás revisar si el elemento ya existe en Wikidata y seleccionarlo, con la casilla marcada o “Crear un elemento nuevo”.

Una vez eliges los elementos en Wikidata que ya existen y los que se deben crear, te aparecerá un listado como el que puede verse en la imagen a continuación. Las flechas rojas señalan los elementos que no existen en Wikidata y se crearán como nuevos.

Puedes usar los filtros para ver únicamente las filas que no se han cotejado. La siguiente imagen muestra la opción para realizar este filtro.

Una vez cotejada la columna de nombres de autor, deberás cotejar la columna de sexo o género. Debes realizar el mismo procedimiento y seleccionar el valor de cada persona.

Para crear una nueva columna con las *Q*'s de cada persona, lo haces mediante la columna de Nombre completo cotejada. Debes seleccionar en la pestaña: **Cotejar > Añadir columna de identificadores de entidades**. Nombra esta columna **Qid**.

Actividad: realiza el cotejamiento de los autores y su género que quieres visibilizar en Wikidata. Marca los que ya existen o declara que los crearás como nuevos.

Posibles problemas en el cotejo de datos Cuando quieres cotejar datos y no aparecen las opciones indicadas en este documento, se pueden buscar clases similares. La siguiente imagen ilustra la búsqueda de una clase mediante el

Carmelo Escoger nueva coincidencia
Josefa <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> María Josefa (67) <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> Crear nuevo elemento Buscar una coincidencia
José Escoger nueva coincidencia
Manuel <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> Manuel (100) <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> Manuel (100) <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> José Manuel (71) <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> Juan Manuel (71) <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> Crear nuevo elemento Buscar una coincidencia

Figure 3: Selección o creación de elementos en Wikidata con OpenRefine

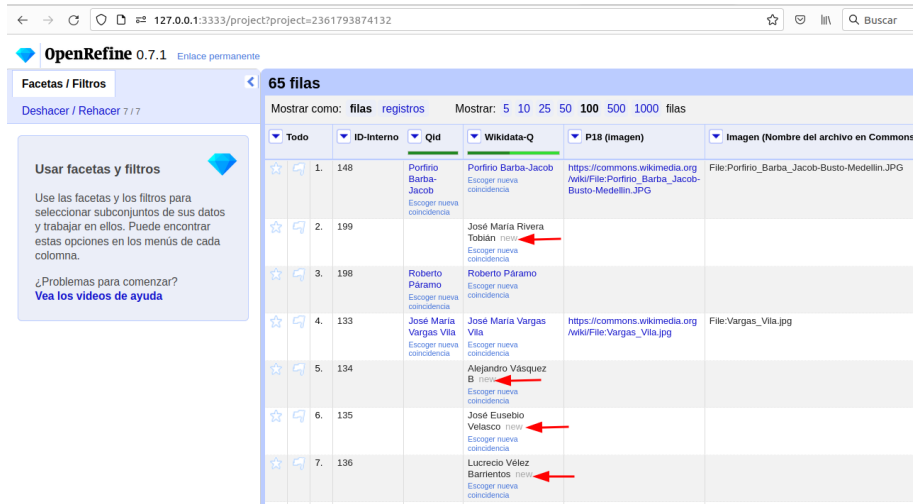


Figure 4: Columna de nombres cotejados

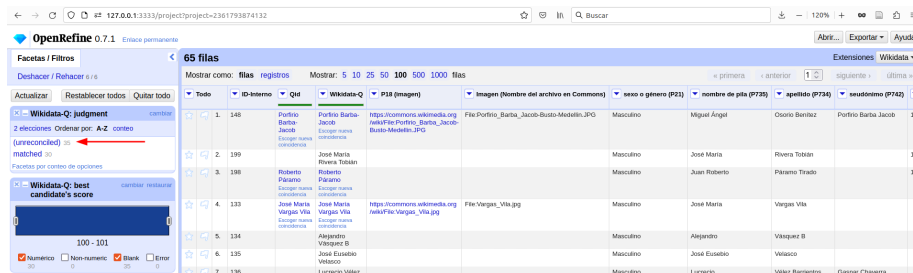


Figure 5: Opción para filtrar los valores no cotejados

cajón de búsqueda. De esta forma, le decimos a OpenRefine exactamente el tipo de dato que estamos buscando.

Cotejar columna "sexo o género (P21)" API del servicio de acceso ↗

Concilia cada celda con una entidad de uno de estos tipos: También use detalles relevantes de otras columnas:

group of living things
Q16334298

biological sex
Q290

grammatical gender
Q162378

columna	¿Incluir?	Como propiedad
ID-Interno	<input type="checkbox"/>	<input type="text"/>
Qid	<input type="checkbox"/>	<input type="text"/>
P31 (instancia de)	<input type="checkbox"/>	<input type="text"/>
Instancia de (Q)	<input type="checkbox"/>	<input type="text"/>
P18 (imagen)	<input type="checkbox"/>	<input type="text"/>
Imagen (Nombre del archivo en Commons)	<input type="checkbox"/>	<input type="text"/>
sexo o género (P21)	<input type="checkbox"/>	<input type="text"/>
Sexo o género (Q)	<input type="checkbox"/>	<input type="text"/>

Cotejar contra la clase:

Cotejar contra ninguna:
male Q6581097

Cotejar automáticamente
to be used in "sex or gender" (P21) to indicate that the human subject is a male or "semantic gender" (P10339) to indicate that a

Máximo número de candida:

Figure 6: Uso del campo “Cotejar contra la clase”.

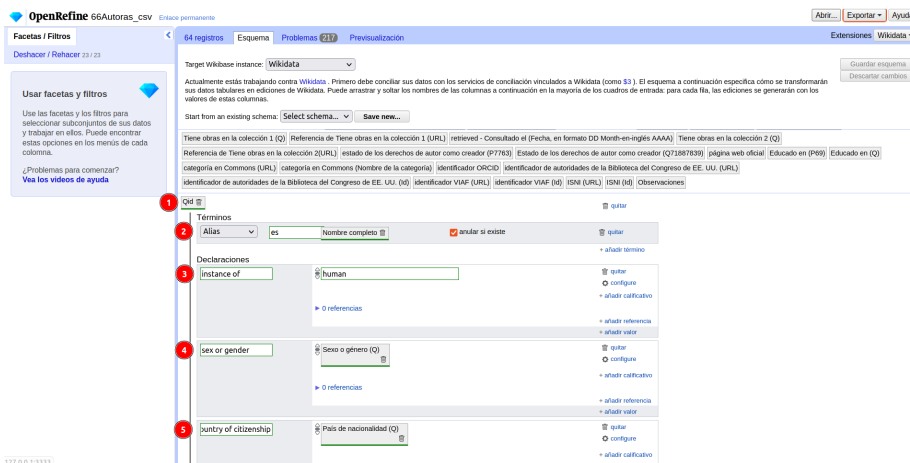
Publicar los datos en Wikidata

Si ya cotejaste los datos de las personas en tu listado, lo que haremos ahora será publicarlos en Wikidata. Para esto, tenemos que crear un *Esquema*. Un esquema es lo que usamos para decirle a Wikidata qué datos actualizaremos y con qué valores.

Crear un esquema y hacer la primera publicación de datos Para crear un esquema, debes ir a la opción superior derecha **Extensiones: Wikidata > Editar esquema de Wikidata**. Allí, deberás seleccionar **Añadir elemento** y empezar a completar los cajones, como deberá verse el resultado en Wikidata. La siguiente imagen lo ilustra.

1. **Qid** tiene el identificador único Q del elemento que se modificará. Si está en blanco, se creará.
2. **Etiqueta, descripción, alias**, el idioma del alias y el nombre completo de la persona. Aparecerá en el primer cuadro del elemento de Wikidata.
3. **Instancia de:** *humano*
4. **Sexo o género:** Según corresponda en la tabla.
5. **País de nacionalidad:** Colombia
6. **Estado de los derechos de autor como creador:** copyrights on works have expired
 - En los valores de esta declaración:
 - **Applies to jurisdiction:** *countries with 80 years pma or shorter*

- En las referencias de esta declaración:
 - **Based on heuristic: *died more than 80 years ago***



Puedes ir a la pestaña de previsualización para revisar cómo quedarán los datos una vez cargados a Wikidata. En la siguiente imagen, puedes ver un ejemplo.



Figure 7: Previsualización de carga de datos

En los casos en los que se quiera incluir nuevos elementos a Wikidata, se deberá modificar en el esquema los datos de la columna que se encuentren vacíos. Como ejemplo, en la siguiente imagen se incluyen los datos de la columna “Nombre

completo”.

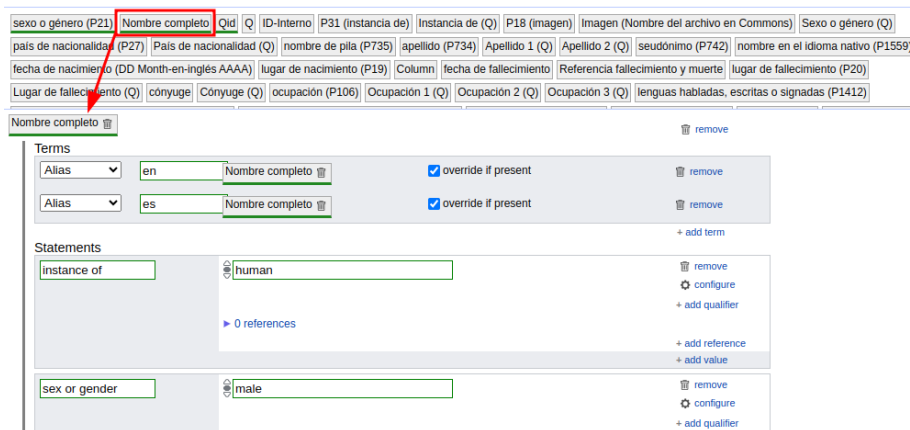


Figure 8: Crear elementos a partir de Nombre Completo

Recuerda que en esta carga de datos, lo que buscamos es que se identifiquen las obras que se encuentran en dominio público. Ten cuidado porque debes subir los datos de derechos de autor con valores de referencia. En la siguiente imagen puedes ver los valores y las referencias.



Figure 9: La delcaración de una obra en dominio público con su referencia.

En **Extensiones: Wikidata > Cargar los datos a Wikidata** podrás publicar lo que tienes en OpenRefine a Wikidata. Deberás escribir el mensaje que describe la actualización que haces.

Otra forma de actualizar los datos es mediante **Exportar > Archivo de QuickStatements**. Esto genera un documento de texto. Puedes abrirlo con un editor y copiar su contenido, para luego pegarlo en el sitio de QuickStatements. Deberás darle **Lote nuevo** y pegar el texto en cajón. Podrás escribir el **nombre de lote** e **importar órdenes de V1**.

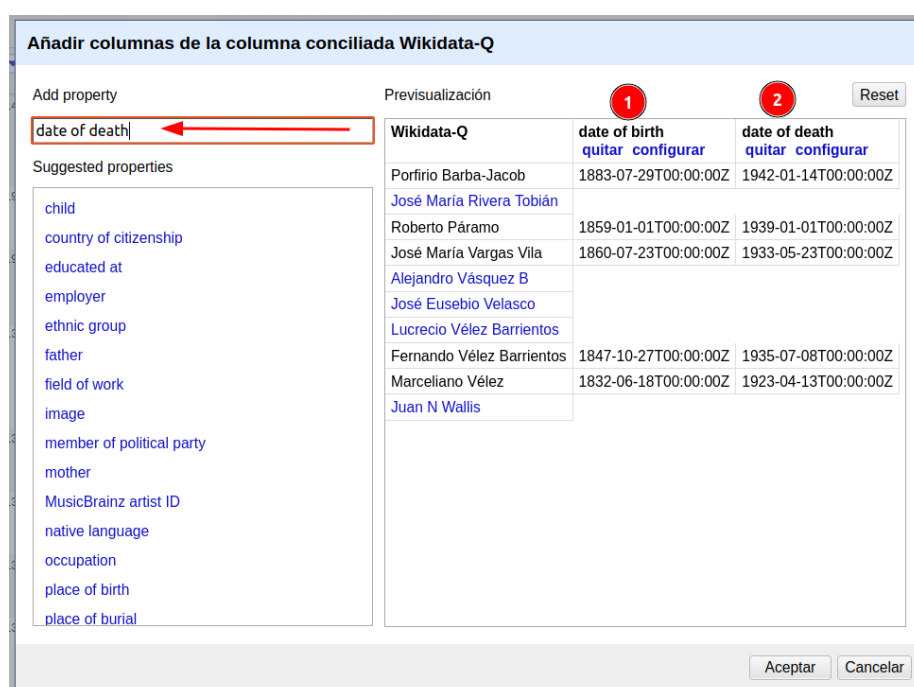
Nota: Una vez organizados los datos, si se carga por quickstatement, incluir la nota de actualización. Por ejemplo: “personas en dominio público para Colombia”.

Actividad: carga en Wikidata las personas de tu set de datos con las 7 variables mencionadas en esta sección.

Publicar las fechas de nacimiento y muerte

El paso siguiente es actualizar fechas de nacimiento y muerte de las personas. Si es posible, identificadores como el VIAF. Para hacerlo, extraigo los datos de Wikidata a partir de las Q's. Así, veo qué tiene Wikidata y qué necesita completarse.

La ruta es: Wikidata-Q > Editar columna > Añadir columnas de valores cotejados. De esta manera, puedo agregar las columnas de años. Como las propiedades no existen, deben digitarse. Al seleccionarlas, aparecen. Repite este proceso por cada columna que quieras incluir en Wikidata. La imagen lo muestra:



The screenshot shows the Wikidata interface for adding columns to a reconciled column. The 'Add property' field contains 'date of death'. The 'Previsualización' table shows columns for 'date of birth' and 'date of death' with corresponding Wikidata-Q values and dates.

Wikidata-Q	date of birth quitar configurar	date of death quitar configurar
Porfirio Barba-Jacob	1883-07-29T00:00:00Z	1942-01-14T00:00:00Z
José María Rivera Tobían		
Roberto Páramo	1859-01-01T00:00:00Z	1939-01-01T00:00:00Z
José María Vargas Vila	1860-07-23T00:00:00Z	1933-05-23T00:00:00Z
Alejandro Vásquez B		
José Eusebio Velasco		
Lucrecio Vélez Barrientos		
Fernando Vélez Barrientos	1847-10-27T00:00:00Z	1935-07-08T00:00:00Z
Marceliano Vélez	1832-06-18T00:00:00Z	1923-04-13T00:00:00Z
Juan N Wallis		

Figure 10: Añadir columnas de valores conciliados con fechas de nacimiento y muerte.

Salen las columnas con datos en color verde. Wikidata descarga diferentes tipos de datos dependiendo de si sólo tiene el año o incluso el día. La siguiente imagen lo muestra.

En la imagen, en el 1) está el año, mes y día. En un campo vacío como puede verse en 2) en la imagen, debemos completar nosotros los datos, pero únicamente con el año en formato YYYY. En 3) Wikidata nos muestra los datos que importa cuando tiene únicamente el año.

Ya que los valores de fechas requieren referencias, podemos usar las URLs de donde tomamos las fechas para incluirlas como referencias de los valores que estamos actualizando. Creamos una columna que se llame "Referencia-fecha". Al completar los datos queda como muestra la siguiente imagen.

Wikidata-Q	date of birth	date of death	Qs
Alejandro Vedit	1887-06-24T00:00:00Z	1943-02-09T00:00:00Z	Q42727696
Milciades Garavito Sierra			Q123985148
Antonio José Iregui	1864-01-17T00:00:00Z	1941-01-01T00:00:00Z	Q85865287
Augusto Ordóñez Moreno			Q123985149
Manuel José Sierra Ríos	1885-01-01T00:00:00Z	1941-01-01T00:00:00Z	Q94515008
Heraclio Uribe Uribe			Q123985150
Tomás Vargas	1908-10-23T00:00:00Z	1941-12-21T00:00:00Z	Q23498603
Juan Francisco Aguilera			Q123985151

Figure 11: Imagen con datos 1) descargados de Wikidata y 2) Sin datos en Wikidata.

65 registros						
Mostrar como: filas registros		Mostrar: 5 10 25 50 100 500 1000 registros				
▼ Todo	▼ ID-Interno	▼ Qid	▼ Qs	▼ date of birth	▼ date of death	▼ Referencia-fecha
☆	1.	148	Porfirio Barba-Jacob <small>Escoger nueva coincidencia</small>	Q1744460		
☆	2.	199		Q123985129	1836	1939 1
☆	3.	198	Roberto Páramo <small>Escoger nueva coincidencia</small>	Q109858607		
☆	4.	133	José María Vargas Vila <small>Escoger nueva coincidencia</small>	Q622861	1860-07-23T00:00:00Z	1933-05-23T00:00:00Z 2
☆	5.	134		Q123985130	1881	1928
☆	6.	135		Q123985131	1888	1928

Figure 12: Actualización de fechas con referencias

En la anterior imagen, se ve en 1) las fechas que incluyo para actualizar con referencias. En 2) las fechas que Wikidata ya tiene y por lo tanto no debo actualizar. No es necesario ponerle referencias si no las vamos a actualizar.

Luego, creamos el esquema de datos para actualizar las fechas. Se ve en la siguiente imagen.

Para verificar, puedo ir a la previsualización. Se ve como se muestra en la siguiente imagen. Después, cargamos los datos a Wikidata.

Los nombres

Wikidata debe contener los nombres compuestos como un único valor. Es decir, si una persona se llama Miguel Ángel, como es el caso de Porfirio Barba Jacob, en el campo “Nombre” debes incluir el nombre compuesto “Miguel Ángel” (Q40440281) y no “Miguel” y aparte “Ángel”. Por tal motivo repetimos el proceso: cotejamos los datos y buscamos los nombres compuestos para incluirlos.

Cuando el nombre compuesto no existe, puedes crear elementos nuevos buscando coincidencias. La siguiente imagen muestra cómo hacerlo. En el cajón de búsqueda escribes el nombre (aparece en 1) y si no se encuentra, puedes crearlo como elemento nuevo (aparece en 2).

Después de haber cotejado todos los nombres, puedes volver a editar el esquema de datos para agregar los nombres. La siguiente imagen muestra cómo se ve el esquema de datos con los nombres.

Al ver la previsualización de los datos, puedes ver cómo quedará el resultado

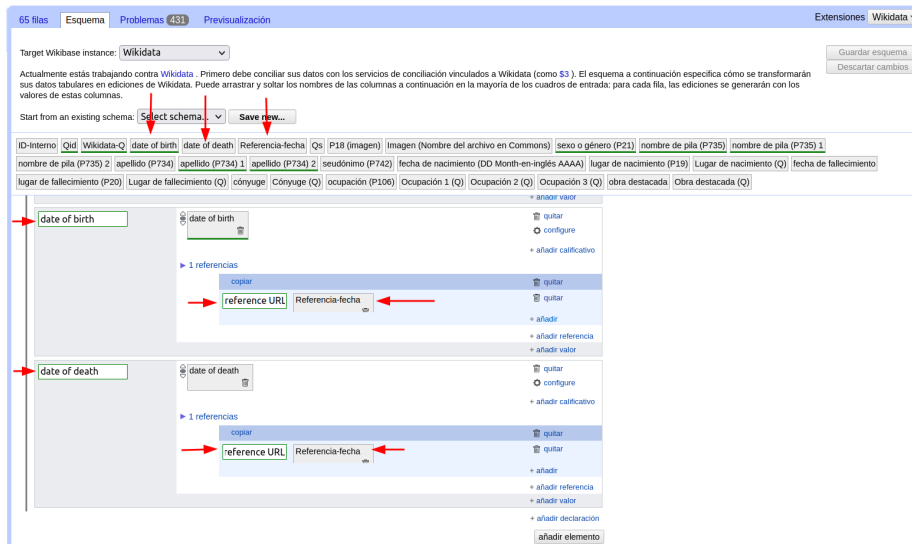


Figure 13: Esquema para actualizar las fechas de nacimiento y muerte.

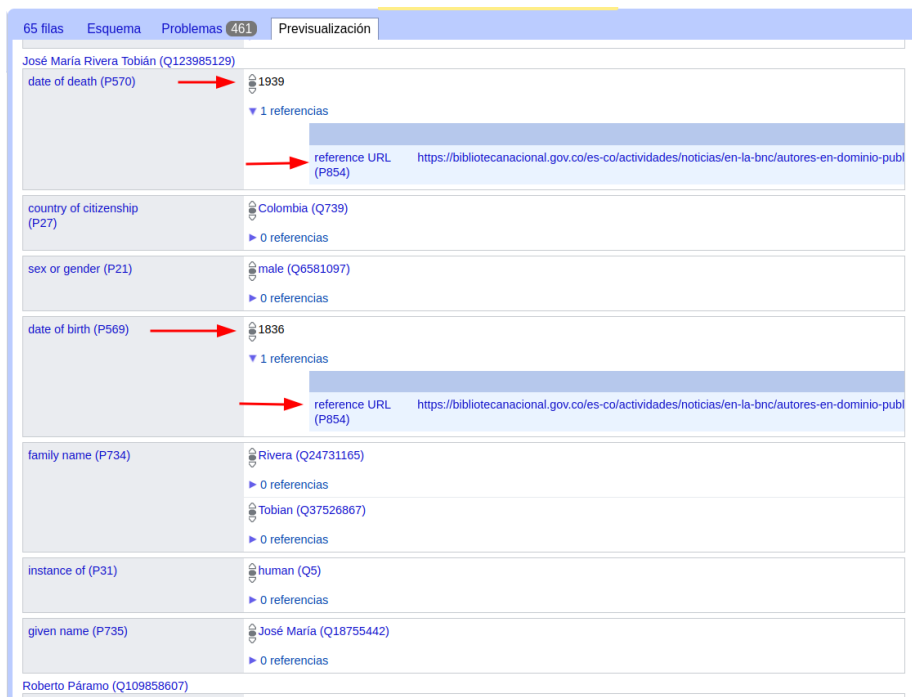


Figure 14: Previsualización de fechas de nacimiento y muerte.

nombre de pila (P735)

Edgar Poe new
[Escoger nueva coincidencia](#)

Melitón
[Escoger nueva coincidencia](#)

Tiberio de Jesús new
[Escoger nueva coincidencia](#)

Gregorio
[Escoger nueva coincidencia](#)

Francisco Cristóbal
 Crear nuevo elemento
[Buscar una coincidencia](#)

Antonio José
[Escoger nueva coincidencia](#)

Luis Aurelio
 Crear nuevo elemento
[Buscar una coincidencia](#)

Alejandro
[Escoger nueva coincidencia](#)

Milciades
 Crear nuevo elemento
[Buscar una coincidencia](#)

Antonio José
[Escoger nueva coincidencia](#)

Figure 15: Cotejo de nombres

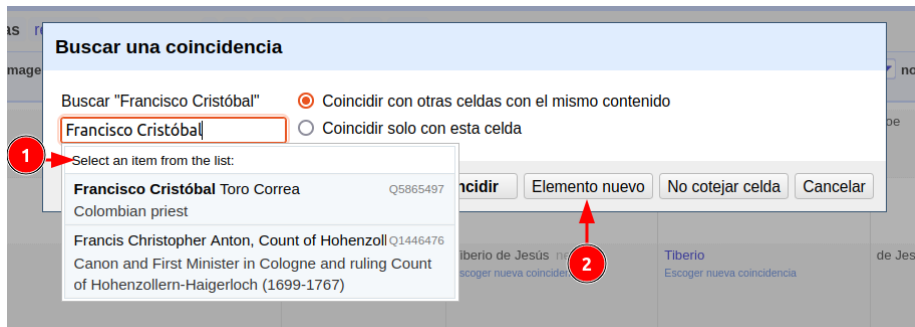


Figure 16: Creación de nombres compuestos cuando no existen

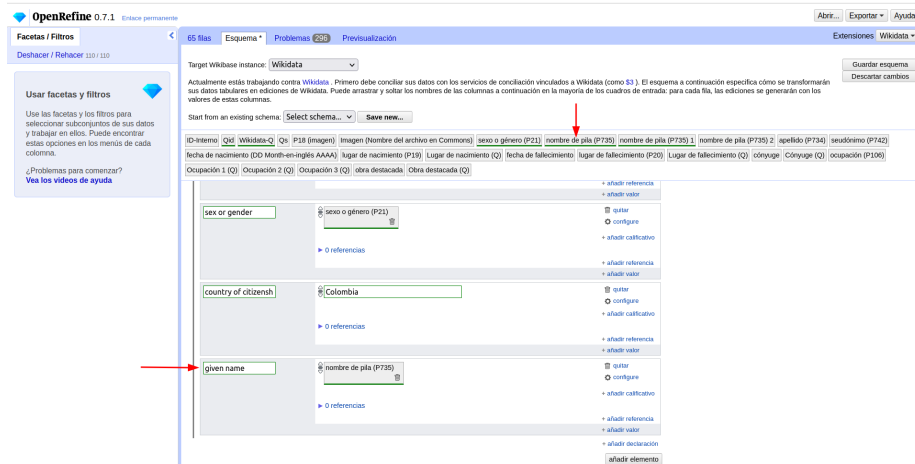


Figure 17: Actualización del esquema de datos para incluir el nombre de pila

para los nombres compuestos.

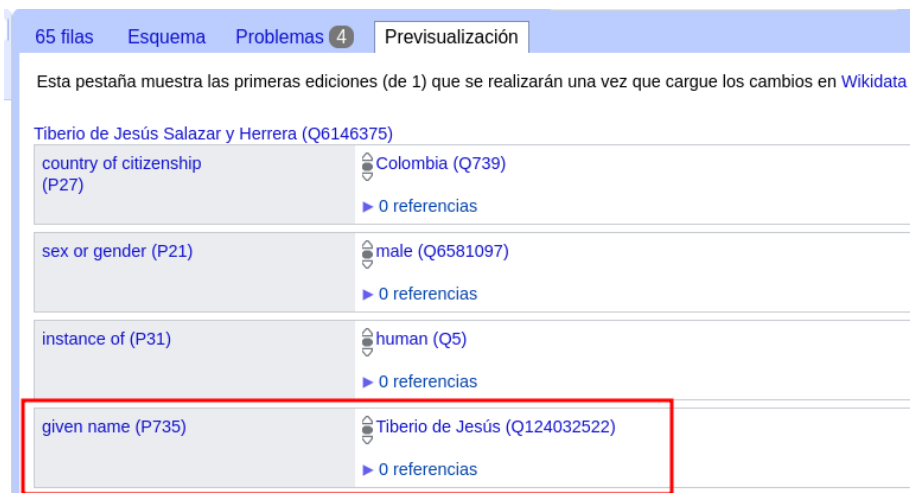


Figure 18: Previsualización de la actualización de nombres

Los apellidos

Para cotejar la columna de los primeros apellidos, sigue los mismos pasos que hemos venido realizando. Coteja los datos con apellidos.

Recuerda que el primer apellido va en la propiedad P734 y el segundo apellido en la propiedad P1950. Puedes revisar en nuestro ejemplo, el de Porfirio Barba Jacob.

Los identificadores

Si bien no son campos obligatorios debido a que no todas las personas en dominio público tienen identificadores, recomendamos que se diligencien. Aquí pondremos el ejemplo del diligenciamiento del identificador VIAF. De esta forma, las personas tendrán visibilidad en otras instituciones que las comparan con sus registros para establecer equivalencias.

Crea la columna para traer los valores que Wikidata tenga de VIAF. La ruta es: Wikidata-Q > Editar columna > Añadir columnas de valores cotejados. Ya que VIAF no aparece entre los primeros resultados, deberás introducir su nombre para que aparezca.

Esto creará una nueva columna. Algunos campos no tienen datos de su identificador VIAF. Esto quiere decir que deberás ir a buscarlos al sitio web de VIAF.

Una vez completas los datos en OpenRefine de VIAF, puedes actualizar el esquema de datos para cargar los archivos en Wikidata. La siguiente imagen muestra cómo queda el nuevo esquema de datos con el identificador VIAF. Si encuentras varios identificadores VIAF para una misma persona, te recomendamos que los incluyas todos.

Cotejar columna "apellido (P734) 1"

Concilie cada celda con una entidad de uno de estos tipos:

- Q16521 department of Uruguay Q56059
- language Q34770
- municipality of Galicia Q2276925
- literary work Q7725634
- municipality of Colombia Q2555896
- family name Q101352
- female given name Q11879590
- city in the United States Q1093829

También use detalles relevantes de otras columnas:

columna	¿Incluir? Como propiedad
ID-Interno	<input type="checkbox"/>
Qid	<input type="checkbox"/>
Wikidata-Q	<input type="checkbox"/>
Qs	<input type="checkbox"/>
P18 (imagen)	<input type="checkbox"/>
Imagen (Nombre del archivo en Commons)	<input type="checkbox"/>
sexo o género (P21)	<input type="checkbox"/>
nombre de pila (P735)	<input type="checkbox"/>
nombre de pila (P735) 1	<input type="checkbox"/>

Cotejar contra la clase:

Cotejar contra ninguna clase en particular

Cotejar automáticamente candidatos con alta confianza

Máximo número de candidatos a devolver

Añadir servicio estándar... Discover services... Empezar a cotejar... Cancelar

Figure 19: Cotejo de datos con apellidos

Añadir columnas de la columna conciliada Wikidata-Q

Add property:

Previsualización

Wikidata-Q	VIAF ID
Porfirio Barba-Jacob	17237110
José María Rivera Tobión	
Roberto Páramo	16296197
José María Vargas Vila	39419512
Alejandro Vásquez B	
José Eusebio Velasco	
Lucrecio Vélez Barrientos	
Fernando Vélez Barrientos	
Marceliano Vélez	
Juan N Wallis	
Liborio Zerda	31539238
Luis Zea Uribe	
Pedro Antonio Zubieta	
Eduardo Zuleta Gaviria	
Carmelo Arango Martínez	
Josefa Andrade Berti	
José Antonio Archila	
Manuel Arteaga Hernández	44381832

Suggested properties:

- child
- country of citizenship
- educated at
- employer
- ethnic group
- father
- field of work
- image
- member of political party
- mother
- MusicBrainz artist ID
- native language
- occupation
- place of birth
- place of burial

Aceptar Cancelar

Figure 20: Columna de VIAF ID

65 registros						
Esquema Problemas 440 Previsualización						
Mostrar como: filas registros						
Mostrar: 5 10 25 50 100 500 1000 registros						
▼ Todo	▼ ID-Interno	▼ Qid	▼ Wikidata-Q	▼ VIAF ID	▼ Qs	
☆	1.	148	Porfirio Barba-Jacob Escoger nueva coincidencia	Porfirio Barba-Jacob Escoger nueva coincidencia	17237110	Q1744460
☆	2.	199		José María Rivera Tobión Escoger nueva coincidencia		Q123985129
☆	3.	198	Roberto Páramo Escoger nueva coincidencia	Roberto Páramo Escoger nueva coincidencia	16296197	Q109858607
☆	4.	133	José María Vargas Vila Escoger nueva coincidencia	José María Vargas Vila Escoger nueva coincidencia	39419512	Q622861

Figure 21: Personas sin VIAF en OpenRefine

65 registros Esquema Problemas 440 Previsualización

Target Wikibase instance: Wikidata

Actualmente estás trabajando contra Wikidata. Primero debe conciliar sus datos con los servicios de conciliación vinculados a Wikidata (como \$3). El esquema a co sus datos tabulares en ediciones de Wikidata. Puede arrastrar y soltar los nombres de las columnas a continuación en la mayoría de los cuadros de entrada: para ca valores de estas columnas.

Start from an existing schema: select schema... Save new...

ID-Interno	Qid	Wikidata-Q	VIAF ID	Qs	date of birth	date of death	Referencia-fecha	P18 (imagen)	Imagen (Nombre del archivo en Commons)	sexo o género					
nombre de pila (P735) 2	apellido (P734) 2	apellido (P734) 1	apellido (P734) 2	seudónimo (P742)	fecha de nacimiento (DD Month-en-inglés AAAA)	lugar de naci	lugar de fallecimiento (P20)	Lugar de fallecimiento (Q)	cónyuge (Q)	Cónyuge (Q)	ocupación (P106)	Ocupación 1 (Q)	Ocupación 2 (Q)	Ocupación 3 (Q)	obra des

date of death

Spanish

VIAF ID

+ añadir referencia
+ añadir valor
quitar
configure
añadir calificativo
1 referencias
añadir referencia
añadir valor
añadir referencia
añadir valor
añadir referencia
añadir valor
añadir declaración
añadir elemento

Figure 22: Esquema de datos con el identificador VIAF

Ya que los identificadores VIAF necesitan referencias, puedes incluir esta estructura de referencias. Declaras que lo encontraste en VIAF, indicas el identificador y la fecha en que lo viste:

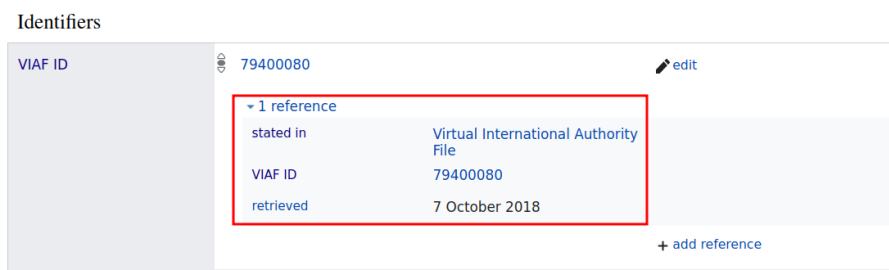


Figure 23: Referencias de un identificador VIAF

Al actualizar el esquema de datos, quedará con los datos de referencias en VIAF. La siguiente imagen lo ilustra.

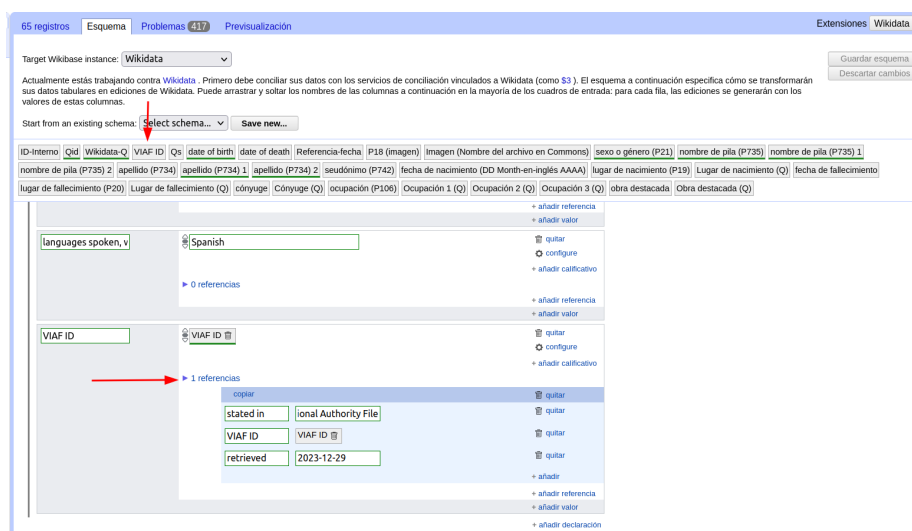


Figure 24: Esquema de datos de VIAF con referencias.

Ocupaciones

En este caso, no traeremos los datos de Wikidata, sino que usaremos las columnas de nuestra hoja de cálculo. Como la columna de “Ocupaciones” se puede repetir cuantas veces lo necesites, pueden haber múltiples columnas.

Cotejamos las diferentes columnas de Ocupaciones. En nuestro ejemplo, tenemos hasta 3 ocupaciones por persona. Cotejamos cada una de las columnas de Ocupaciones. Wikidata nos ofrece “Ocupaciones” para identificar las que ya tiene registradas. La siguiente imagen lo muestra.

Se obtendrán varias celdas con los valores cotejados y algunas en blanco, que se podrán diligenciar en Wikidata.

Cotejar columna "Ocupación 1 (Q)" API del servicio de acceso

Concilie cada celda con una entidad de uno de estos tipos:

artistic profession
 Q88789639
 occupation ←
 Q12737077

También use detalles relevantes de otras columnas:

columna	¿Incluir? Como propiedad
ID-Interno	<input type="checkbox"/>
Qid	<input type="checkbox"/>
VIAF ID	<input type="checkbox"/>
Qs	<input type="checkbox"/>
date of birth	<input type="checkbox"/>
date of death	<input type="checkbox"/>
P18 (imagen)	<input type="checkbox"/>
Imagen (Nombre del archivo en Commons)	<input type="checkbox"/>
sexo o género (P21)	<input type="checkbox"/>
nombre de pila (P735)	<input type="checkbox"/>

Cotejar contra la clase:
 Cotejar contra ninguna clase en particular
 Cotejar automáticamente candidatos con alta confianza
 Máximo número de candidatos a devolver

Figure 25: Cotejo de Ocupaciones

500 1000 registros « primera < anterior 1 siguiente > última »

Ocupación 1 (Q)	Ocupación 2 (Q)	Ocupación 3 (Q)	Wikidata-Q	obra destacada	Obra destacada (Q)	Referencia-fecha
			Porfirio Barba-Jacob Escoger nueva coincidencia	La canción de la vida profunda y otros poemas	Q123029086	
			José María Rivera Toblián Escoger nueva coincidencia			https://bibliotecanacional.gov.co/es-co/actividades/noticias/en-la-bnc/autores-en-dominio-publico-2020
painter Escoger nueva coincidencia	artist Escoger nueva coincidencia		Roberto Páramo Escoger nueva coincidencia	Roberto Páramo: pintor de la sabana		
			José María Vargas Vía Escoger nueva coincidencia			
			Alejandro Vásquez B			https://catalogo.bibliotecanacional.gov.co/

Figure 26: Ocupaciones. Algunas cotejadas, otras por diligenciar.

En nuestro ejemplo, aparecen diferentes opciones para seleccionar la ocupación que mejor describe la que tenemos para nuestras personas.

Ocupación 1 (Q)	Ocupación 2 (Q)	Ocupación 3 (Q)	Wikidata-Q	obra destacada	Obra destacada (Q)	Referencia-fecha
diplomat Escoger nueva coincidencia	Escritor <input checked="" type="checkbox"/> writer (100) ← <input checked="" type="checkbox"/> author (100) <input checked="" type="checkbox"/> physician writer (73) <input checked="" type="checkbox"/> screenwriter (67) <input checked="" type="checkbox"/> playwright (67) <input checked="" type="checkbox"/> travel writer (67) <input checked="" type="checkbox"/> short story writer (67) <input checked="" type="checkbox"/> ghostwriter (64) <input checked="" type="checkbox"/> opinion journalist (64) <input checked="" type="checkbox"/> comics writer (64) <input checked="" type="checkbox"/> hymnwriter (64) <input checked="" type="checkbox"/> correspondent (62) <input checked="" type="checkbox"/> non-fiction writer (62) <input checked="" type="checkbox"/> dialect writer (62) <input checked="" type="checkbox"/> novelist (56) <input checked="" type="checkbox"/> crime fiction writer (56) <input checked="" type="checkbox"/> children's writer (55) <input checked="" type="checkbox"/> science fiction writer (53) <input checked="" type="checkbox"/> Crear nuevo elemento Buscar una coincidencia		Juan N Wallis Escoger nueva coincidencia			https://web.archive.org/web/20170211105046/https://www.banrepublicanal.org/blaa/literatura/quien/quien23.htm
physician Escoger nueva coincidencia	scientist Escoger nueva coincidencia		Liborio Zerda Escoger nueva coincidencia			https://enciclopedia.banrepublica.org/index.php?title=Liborio_Zerda
physician Escoger nueva coincidencia	essayist Escoger nueva coincidencia	Académico	Luis Zea Uribe Escoger nueva coincidencia			https://web.archive.org/web/20190412120516/http://pasioncreadora.info/ensayo/la-literatura-de-artioquia/ensayos-nombres-y-obras/
			Pedro Antonio Zubieta Escoger nueva coincidencia			https://catalogo.bibliotecanacional.gov.co/

Figure 27: Cotejando ocupaciones en OpenRefine

Luego de realizar todo el cotejamiento, procedemos a actualizar el esquema de datos.

Verificamos que los datos hayan quedado bien en la previsualización. Si el resultado es el esperado, podemos proceder a cargar los datos.

Como podrás ver, estamos haciendo un proceso muy similar una y otra vez. Ajustamos los datos en OpenRefine, pero como son datos muy parecidos podemos actualizar a todas las personas que estén en dominio público de forma muy rápida. Repite este proceso cuantas veces sea necesario.

Las obras destacadas

Una vez creadas las personas en Wikidata, nos interesa enlazar los datos a las obras que hayan realizado, que son las que están en dominio público. En nuestro caso, algunas personas tenían obras destacadas en el listado que la Biblioteca Nacional de Colombia proponía.

En este punto hay dos tareas a realizar: por una parte, repetir el proceso de creación de elementos, ya no para las personas, sino para sus obras. En términos generales, hacemos lo mismo: verificamos si existe en Wikidata. Si existe, añadimos los datos que podamos aportar. Si no existe, creamos el nuevo elemento. Pero la segunda parte es la que puede jugar nos malas pasadas. ¿Cómo definimos cuál es la obra u obras destacadas?

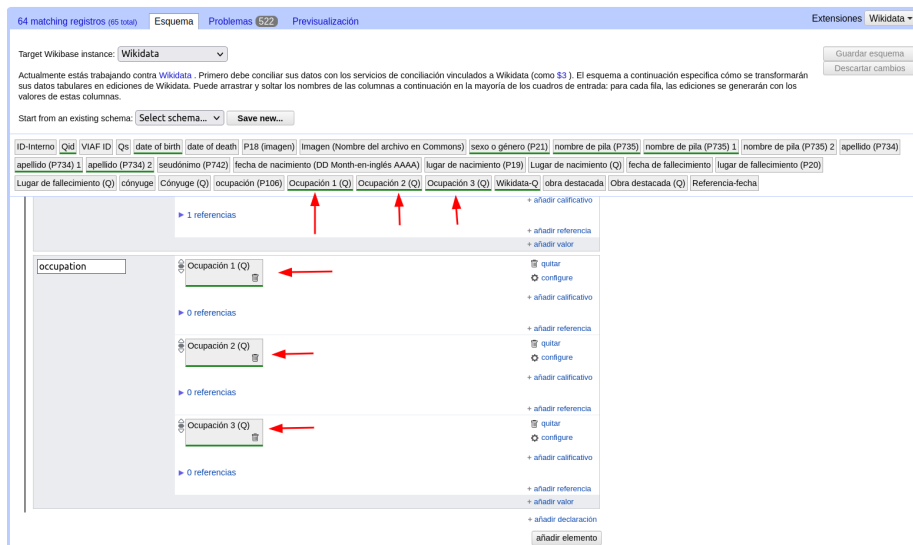


Figure 28: Esquema de datos de las ocupaciones

Proponemos que una forma de identificarla sea a través de identificadores como el VIAF. Es decir, si múltiples instituciones tienen la obra en sus bibliotecas, esto quiere decir que la obra es más popular.

Una vez definida la obra, se procede a hacer dos cosas: al elemento “Obra” se le asigna una autora, que fue lo que hicimos anteriormente. De esta forma garantizamos que todas las obras tienen una persona relacionada que existe en Wikidata. Por otra parte, debemos ir al perfil de las personas para asignarle una obra destacada.

Si bien el concepto de “Obra destacada” puede ser subjetivo y discutible, proponemos que el hecho de que una obra haya llegado al catálogo de una biblioteca sea un parámetro para considerarlo notable. Es decir, al menos pasó por el criterio de quien realizó el proceso de selección y adquisiciones en una institución.

No repetiremos cada uno de los pasos en OpenRefine, pero aquí tienes cómo se vería el esquema de datos para las obras. En este ejemplo, ya hemos realizado el cotejamiento y limpieza de todos los datos, verificando que los elementos existen en Wikidata o se deben crear como nuevos.

Ya que los tipos de obras pueden variar, revisa la columna de tipos de obra. Generalmente encontrarás obras escritas. Sin embargo, también puedes tener pinturas o música. En la siguiente imagen verás el ejemplo de los tipos de obra y los tipos de autor. Los músicos son “compositores”.

Si las fechas de publicación no son exactas, puedes hacer que OpenRefine publique en Wikidata información con precisión de décadas. Para esto, debes agregar la letra “D” al final del año de publicación de la obra. En la Guabina santandereana se puede comprobar que se refiere a una década. En este enlace se encuentra el manual que lo explica.

Por el lado de las personas, así queda el esquema de datos para incluir su obra

64 matching registros (65 total) Esquema Problemas 522 Previsualización

▶ 0 referencias

Lucrecio Vélez Barrientos (Q123985132)

VIAF ID (P214) 21169375640224282977
▶ 1 referencias

date of death (P570) 1925
▶ 1 referencias

occupation (P106) novelist (Q6625963) ←

country of citizenship (P27) Colombia (Q739)
▶ 0 referencias

sex or gender (P21) male (Q6581097)
▶ 0 referencias

languages spoken, written or signed (P1412) Spanish (Q1321)
▶ 0 referencias

date of birth (P569) 1850
▶ 1 referencias

instance of (P31) human (Q5)
▶ 0 referencias

given name (P735) Lucrecio (Q124032490)
▶ 0 referencias

Figure 29: Previsualización de ocupaciones

https://viaf.org/viaf/39419512/

100 1_1a Vargas Vila, José María, 1850-1933

200_1a Vargas Vila, J. M., 1850-1933

400s: Formas alternativas del nombre (56)

500s: Nombres relacionados (2)

551_1a Barcelona 14 ors 14 https://id-nb.info/standards/elementset/grd/placeOfDeath

551_1a Bogotá 14 ors 14 https://id-nb.info/standards/elementset/grd/placeOfBirth

Works

Show 10 entries

Title

Minotaruo . novela

Novena sinfonia . novela

Works_1920

Works_1972

Rosas de la tarde

Works_Selecciones

Works_1920 Siemra

Works_1998

Works_1916 Antonio Rubiñes

Works_1930

Sources

Showing 1 to 10 of 89 entries

Previous 1 2 3 4 5 ... 9 Next

Figure 30: Obras de José María Vargas Vila en VIAF, con “Minotaruo” y “Rosa de la tarde” en múltiples instituciones alrededor del mundo

65 filas **Esquema** Problemas **289** Previsualización

Target Wikibase instance: Wikidata

Actualmente estás trabajando contra Wikidata . Primero debe conciliar sus datos con los servicios de conciliación vinculados a Wikidata (como \$3). El esquema sus datos tabulares en ediciones de Wikidata. Puede arrastrar y soltar los nombres de las columnas a continuación en la mayoría de los cuadros de entrada; por valores de estas columnas.

Start from an existing schema: **Select schema...** **Save new...**

ID-Interno	Qid	título (P1476)	autor (P50)	ID-Interno-Autor	Referencia	Observaciones	Tipo de obra																																																																					
título (P1476) <div style="float: right;">🗑️ quitar</div> <div> <p>Términos</p> <table border="1"> <tr> <td>Descripción</td> <td>es</td> <td>Obra colombiana</td> <td><input type="checkbox"/> anular si existe</td> <td>🗑️ quitar</td> </tr> <tr> <td>Etiqueta</td> <td>es</td> <td>título (P1476)</td> <td><input type="checkbox"/> anular si existe</td> <td>🗑️ quitar</td> </tr> </table> <p style="text-align: right;">+ añadir término</p> </div> <div> <p>Declaraciones</p> <table border="1"> <tr> <td>instance of</td> <td>Tipo de obra</td> <td>🗑️ quitar</td> <td>⚙️ configure</td> <td>+ añadir calificativo</td> </tr> <tr> <td colspan="5">▶ 0 referencias</td> </tr> <tr> <td colspan="5">+ añadir referencia</td> </tr> <tr> <td colspan="5">+ añadir valor</td> </tr> </table> <table border="1"> <tr> <td>title</td> <td>es</td> <td>título (P1476)</td> <td>🗑️ quitar</td> <td>⚙️ configure</td> <td>+ añadir calificativo</td> </tr> <tr> <td colspan="6">▶ 1 referencias</td> </tr> <tr> <td colspan="6">copiar</td> </tr> <tr> <td>reference URL</td> <td>Referencia</td> <td>🗑️ quitar</td> <td>🗑️ quitar</td> <td>+ añadir</td> </tr> <tr> <td colspan="6">+ añadir referencia</td> </tr> <tr> <td colspan="6">+ añadir valor</td> </tr> </table> <table border="1"> <tr> <td>author</td> <td>autor (P50)</td> <td>🗑️ quitar</td> <td>⚙️ configure</td> </tr> </table> </div>								Descripción	es	Obra colombiana	<input type="checkbox"/> anular si existe	🗑️ quitar	Etiqueta	es	título (P1476)	<input type="checkbox"/> anular si existe	🗑️ quitar	instance of	Tipo de obra	🗑️ quitar	⚙️ configure	+ añadir calificativo	▶ 0 referencias					+ añadir referencia					+ añadir valor					title	es	título (P1476)	🗑️ quitar	⚙️ configure	+ añadir calificativo	▶ 1 referencias						copiar						reference URL	Referencia	🗑️ quitar	🗑️ quitar	+ añadir	+ añadir referencia						+ añadir valor						author	autor (P50)	🗑️ quitar	⚙️ configure
Descripción	es	Obra colombiana	<input type="checkbox"/> anular si existe	🗑️ quitar																																																																								
Etiqueta	es	título (P1476)	<input type="checkbox"/> anular si existe	🗑️ quitar																																																																								
instance of	Tipo de obra	🗑️ quitar	⚙️ configure	+ añadir calificativo																																																																								
▶ 0 referencias																																																																												
+ añadir referencia																																																																												
+ añadir valor																																																																												
title	es	título (P1476)	🗑️ quitar	⚙️ configure	+ añadir calificativo																																																																							
▶ 1 referencias																																																																												
copiar																																																																												
reference URL	Referencia	🗑️ quitar	🗑️ quitar	+ añadir																																																																								
+ añadir referencia																																																																												
+ añadir valor																																																																												
author	autor (P50)	🗑️ quitar	⚙️ configure																																																																									

Figure 31: Esquema de datos de obras

▼ Tipo de obra	▼ Tipo de autor
Escoger nueva coincidencia	
written work Escoger nueva coincidencia	author
written work Escoger nueva coincidencia	author
written work Escoger nueva coincidencia	author
written work Escoger nueva coincidencia	author
musical work/composition Escoger nueva coincidencia	composer
written work Escoger nueva coincidencia	author

Figure 32: Tipos de autores y tipos de obras

as 5 Previsualización			
Mostrar: 5 10 25 50 100 500 1000 filas			
▼ título (P1476)	▼ Fecha publicación	▼ Referencia-fecha	▼ autor (P50)
Cantas del Valle de Tenza Escoger nueva coincidencia	1949	https://catalogo.bibliotecanacional.gov.co/	Joaquín R. Medina Escoger nueva coincidencia
Guabina santandereana Escoger nueva coincidencia	1910s	https://catalogo.bibliotecanacional.gov.co/	Hernando Lelio Olarte Pardo Escoger nueva coincidencia
Mi negrita: joropo Escoger nueva coincidencia	1952	https://bibliotecanacional.gov.co/es-co/actividades/noticias/en-la-bnc/dia-dominio-publico-2021	Patrocínio Ortiz Escoger nueva coincidencia

Figure 33: Ejemplo de obra con precisión de una década y otras con precisión del año

notable.

Así se ve la previsualización de la inclusión de obras notables.

Tercera parte: difusión y conclusiones

Organizar talleres de difusión

Lo más interesante de todo este proceso es compartir. Por eso te recomendamos que realices talleres de difusión para que más personas hagan este proceso y entre todas se fortalezca el dominio público colombiano.

Los pasos para hacer un taller de difusión son:

1. Ten en cuenta la celebración del día del dominio público. Si bien el 1 de enero de cada año es una fecha en la que la mayoría de personas se encuentra de vacaciones en Colombia, programa un evento en otro momento del año para que la celebración no se olvide.
2. Contacta a aliados, como la Biblioteca Nacional de Colombia, las bibliotecas departamentales o Wikimedia Colombia.
3. Realiza la invitación al evento con suficiente antelación. Entre más diversa sea la participación, mucho mejor.
4. Asegúrate de obtener permisos para que los contenidos creados en el taller tengan licencias libres y que las personas que participan autoricen a aparecer en fotografías o videos que registren la actividad.
5. Documenta la experiencia de quienes participan. Siempre nos vienen las cosas desde afuera, así que mejor si las adaptamos a nuestros contextos y realidades.
6. Publica la información con licencias abiertas.
7. Repite esto el año que viene, con las nuevas personas que entren en dominio público.


65 filas Esquema Problemas 3 Previsualización

Target Wikibase instance: Wikidata

Actualmente estás trabajando contra Wikidata . Primero debe conciliar sus datos con los servicios de conciliación vinculados a Wikidata (como S3). El esquema a c sus datos tabulares en ediciones de Wikidata. Puede arrastrar y soltar los nombres de las columnas a continuación en la mayoría de los cuadros de entrada; para c valores de estas columnas.

Start from an existing schema: Select schema... Save new...

ID-Interno	Qid	título (P1476)	autor (P50)	ID-Interno-Autor	Referencia	Observaciones	Tipo de obra
			autor (P50)				

autor (P50)  quitar

Términos
no se añadió ninguna etiqueta, descripción ni alias + añadir término

Declaraciones

notable work	título (P1476)	quitar	configure	+ añadir calificativo
▶ 1 referencias				
copiar	reference URL	Referencia	quitar	quitar
			+ añadir	+ añadir referencia
			+ añadir valor	+ añadir declaración
añadir elemento				

Figure 34: Esquema de datos de obra notable para las personas

65 filas Esquema Problemas 3 Previsualización

Esta pestaña muestra las primeras ediciones (de 55) que se realizarán una vez que cargue los cambios en Wikidata . Puede utilizar facetas para inspeccionar las ediciones en elementos concretos.

Roberto Páramo (Q109858607)	notable work (P800)	Roberto Páramo: Paisaje, bodegón, ciudad (Q124285374)	▶ 1 referencias
		reference URL	https://bibliotecanacional.gov.co/es-co/actividades/noticias/en-la-bnc/autores-en-dominio-publ (P854)
Liborio Zerda (Q5974856)	notable work (P800)	Dorado (Q124285375)	▶ 0 referencias
Luis Zea Uribe (Q123985134)	notable work (P800)	Producciones escogidas. (Q124285376)	▶ 0 referencias

Figure 35: Previsualización de la actualización de las personas y su campo de obras notables.

Consideraciones finales y conclusiones

Este proceso nos abre al mundo. Wikidata es una herramienta multilingüe y además es uno de los proveedores de datos abiertos enlazados más populares de Internet. Tener visibilidad en Wikidata es tener la oportunidad de llegar a audiencias que por otros medios es muy complicado hacerlo. Además, formar parte del Movimiento Wikimedia implica formar parte de un movimiento que puede ayudar a mejorar lo que hacemos.

El visibilizar el dominio público nos enseña dónde estamos, qué cosas se están mostrando y qué cosas deberían mostrarse más. Tenemos una gran deuda con la obra de las mujeres en el dominio público. Al momento de redactar este documento, a inicios de 2024, Wikidata tenía 225 personas colombianas en dominio público, de las cuales apenas 5 son mujeres. Esto quiere decir que apenas el 2,2% de obras tienen una perspectiva femenina. Y por lo tanto, las colecciones patrimoniales que tenemos están muy desbalanceadas.

Al momento de redactar este documento, de las 225 personas en dominio público que identificamos, 99 tenían artículos en Wikipedia en Español, es decir, apenas el 44% de las personas. Podemos entender que lo público son las obras en dominio público, pero también podemos entender a Wikidata como una base de conocimiento pública, cuyos datos deberíamos cuidar.

Si existen obras que nazcan en dominio público ¿cómo podemos incluirlas en este tipo de actividades? Empezar a debatir sobre este tipo de casos nos abre nuevos panoramas sobre el conocimiento colombiano y cómo lo gestionamos.

El uso de OpenRefine, QuickStatements e incluso Wikidata puede que no sea fácil al inicio. Sin embargo, proponemos aprender en comunidad. Que este proceso nos sirva para empezar a desarrollar habilidades que nos acerquen a la soberanía de nuestros datos y de la tecnología que usamos para manipularlos.

Podemos abrir preguntas interesantes, como por ejemplo que estamos contabilizando el día del dominio público a partir de las personas, pero el conteo del dominio público varía de acuerdo a los formatos de las obras. También, que estamos realizando las celebraciones de dominio público con las personas naturales, pero las personas jurídicas también tienen plazos de protección que vencen y por ahora no se están visibilizando.

Esta es una excelente oportunidad para ejercitar los principios de la ciencia abierta. Podemos compartir lo que sabemos, los datos que producimos, las herramientas que usamos y las metodologías que nos sirvieron. Que nos sirva para adaptar y compartir el conocimiento para quien quiera colaborar a futuro.

Referencias bibliográficas

- Barba-Jacob, P. (n.d.). *La canción de la vida profunda y otros poemas* [Wiki]. Wikidata. Retrieved February 15, 2024, from <https://www.wikidata.org/wiki/Q123029086>
- Biblioteca Nacional de Colombia. (n.d.-a). *Catálogo Autoridades Biblioteca Nacional de Colombia* [Repositorio digital]. Catálogo Autoridades Biblioteca Nacional de Colombia Koha. Retrieved December 13, 2023, from <https://autoridades.bibliotecanacional.gov.co/>

- Biblioteca Nacional de Colombia. (n.d.-b). *Deposito Legal: ¿Qué es?* [Portal]. Biblioteca Nacional de Colombia. Retrieved November 3, 2023, from <https://bibliotecanacional.gov.co/es-co/servicios/profesionales-del-libro/deposito-legal>
- Biblioteca Nacional de Colombia. (2016, January 3). *Colombia celebra el Día del Dominio Público* [Portal]. Biblioteca Nacional de Colombia. <https://web.archive.org/web/20160103225557/http://www.bibliotecanacional.gov.co/content/colombia-celebra-el-dia-del-dominio-publico>
- Hausenblas, M. (2012). *5 estrellas datos abiertos* (E. Béjar, Trans.) [Repositorio digital]. 5 estrellas datos abiertos. <https://5stardata.info/es/>
- Ministerio del Interior y de Justicia, República de Colombia. (2002). *Dirección Nacional de Derecho de Autor* [Portal]. Dirección Nacional de Derecho de Autor. <http://www.derechodeautor.gov.co/htm/HOME.ASP>
- OpenRefine*. (2010). [Portal]. OpenRefine. <https://openrefine.org/>
- Porfirio Barba-Jacob*. (n.d.). [Wiki]. Wikidata. Retrieved February 15, 2024, from <https://www.wikidata.org/wiki/Q1744460>
- QuickStatements*. (n.d.). [Repositorio digital]. Toolforge. Retrieved May 6, 2023, from <https://quickstatements.toolforge.org/#/>
- Ramírez-Ordóñez, D. (2011). *Calculadora de dominio público colombiana* [Sitio web]. Nomono: Laboratorio de ciencia de la información y viandadas. <http://a.hiperterminal.com/dominiopublico>
- Wikimedia Colombia*. (2023, November 11). [Blog]. Wikimedia Colombia. <https://wikimediacolombia.org/>