

Descolonizar los Datos Estructurados de Internet



INFORME

EN OCTUBRE DE 2021, COMO ANTESALA A LA WIKIDATA CON DE ESE AÑO, MÁS DE 40 PARTICIPANTES DE TODO EL MUNDO SE REUNIERON PARA CONVERSAR EN TORNO A DESCOLONIZAR LOS DATOS ESTRUCTURADOS DE INTERNET. ESTE EVENTO FUE ORGANIZADO DE MANERA CONJUNTA POR WHOSE KNOWLEDGE?, WIKI MOVIMIENTO BRASIL Y WIKIMEDIA DEUTSCHLAND. EL PRESENTE DOCUMENTO ES EL INFORME DE AQUEL ENCUENTRO Y LAS REFLEXIONES QUE EMERGIERON EN ÉL.

Imagen por Whose Knowledge?, CC BY-SA 4.0, via [Wikimedia Commons](#)

Imagen por Whose Knowledge?, CC BY-SA 4.0, via [Wikimedia Commons](#)



Contenido

¿Por qué organizamos esta conversación?	4
¿Qué hicimos y cómo lo hicimos?	5
Lo que surgió	6
Perspectivas y provocaciones	6
Imaginaciones e implementaciones	9
Escuchar y aprender	10
¿Qué sigue?	11
Agradecimientos	12

¿Por qué organizamos esta conversación?

Los datos estructurados son la clave del funcionamiento de internet tal y como lo conocemos. Son unidades de información que se organizan de manera que las máquinas puedan leerlas, comprenderlas y procesarlas. Existen innumerables aplicaciones y plataformas construidas sobre tales sistemas de datos estructurados, incluyendo las herramientas de proveedores de contenido como Google e iniciativas colaborativas como Wikidata. Como consecuencia de ello, estos sistemas impactan sobre todos los ámbitos de nuestra vida en línea: de los mapas y los sistemas de navegación, a las críticas cinematográficas y los rankings de películas.

Mediante estos sistemas se gestionan cantidades ingentes de datos, que se ordenan y se clasifican en relación con otros datos. Como consecuencia, la información se estructura y se conforma según unas normas, tradiciones y epistemologías muy concretas. Al

intentar aplicar determinadas categorías al mundo, los datos estructurados prescriben ciertos marcos y cosmovisiones; es decir, se alejan de la con frecuencia supuesta neutralidad de los datos. Si analizamos las dinámicas de poder de los datos estructurados, podemos examinar de quién son las perspectivas, la agenda y las ontologías (categorías de clasificación) en estos sistemas, y quién toma las decisiones para construirlos y sostenerlos. Entonces recién podemos trabajar juntas para construir unos sistemas de datos estructurados más justos e igualitarios.

Esta transformación de los sistemas de datos estructurados — y del conocimiento mismo — requiere de un diálogo constante, así como de un compromiso que no se limite a las empresas tecnológicas o a unos sectores concretos. Se precisa un esfuerzo colectivo de personas particulares, organizaciones y comunidades que posean diferentes formas de conocimiento y experiencia. A esto lo llamamos alianza o «solidaridad en acción». Aunque lo más importante es que este proceso compartido debe

poner en el centro la experiencia, las vivencias y el liderazgo de la mayoría minorizada del mundo.

[Como ya hemos explicado antes](#), apenas una fracción del conocimiento público que se genera en internet versa sobre, o está producido por mujeres, personas racializadas y LGTBIQ, comunidades indígenas o pueblos del Sur global, y está disponible en lenguas que no sean ni el inglés ni otras lenguas coloniales. No obstante, somos la mayoría de la población mundial, la mayoría de las personas que están conectadas a internet y las personas más afectadas por los usos y abusos de los datos estructurados. Somos las comunidades que más sufren las consecuencias de las estructuras de poder y privilegios y los diversos sistemas de opresión, tanto históricos como actuales, que se refuerzan mutuamente: la colonización, el patriarcado, el racismo, la homofobia, el clasismo, la discriminación por castas, el capacitismo y demás. Por eso necesitamos comenzar a descolonizar internet, de forma práctica y no solo en sentido figurado. Necesitamos transformar las diferentes formas de

poder y privilegio que dan forma a las tecnologías digitales que usamos.

Si quieres profundizar acerca de por qué importa la justicia del conocimiento para los datos estructurados, consulta [esta presentación](#) de nuestra codirectora Anasuya Sengupta.

¡Qué presentación tan poderosa y sentida sobre la descolonización del conocimiento y la importancia de los datos estructurados de @Anasuyashh @WhoseKnowledge en #WikidataCon2021. Gracias. No estamos para tapar huecos, sino para cambiar cómo se hacen las cosas.

FLAVIA DORIA
[@AFLAVIADORIA](#) EN [TWITTER](#), 29 DE OCTUBRE DE 2021

Descolonizar los datos estructurados de internet nació de la necesidad acuciante de tener una conversación colectiva, poniendo en el centro las voces que suelen estar marginalizadas. El 13 de octubre de 2021 invitamos a más de 40 activistas, organizadores comunitarios, desarrolladores y otros «aliados inusuales» a un espacio colaborativo, plurilingüe y seguro. Este grupo de poderosos pensadores y hacedores se componía de una mayoría de personas que se identificaban como mujeres (71 %), procedentes o residentes en el Sur global (66 %) y de origen indígena, racializadas o negras (82 %). La conversación sirvió de antesala a la [WikidataCon 2021](#) y fue organizada de manera conjunta por [Whose Knowledge?](#), [Wikimedia Deutschland](#) y [Wiki Movimento Brasil](#). Por tanto, también priorizamos la participación de las comunidades de América Latina y el Caribe.

¿Qué hicimos y cómo lo hicimos?

Descolonizar los datos estructurados de internet se organizó en tres mesas redondas diferentes. En la primera, los ponentes invitados formularon Perspectivas y provocaciones para encuadrar de manera general las conversaciones. A continuación, los participantes se dividieron en siete grupos para debatir sobre Imaginaciones e implementaciones. Finalmente, el público compartió las ideas clave en una sesión de Escuchar y aprender, donde pusimos en común ideas y reflexiones para seguir adelante.

Comenzamos la conversación explicitando tres valores y compromisos clave: amor, respeto y solidaridad. Basándonos en estos principios rectores, quisimos que quienes participaran fueran conscientes de sus posicionamientos y privilegios, que se expresaran tal y como son, de manera completa y multifacética. Nuestro

objetivo fue promover interacciones centradas en la escucha y el aprendizaje, además del respeto (mutuo entre participantes y hacia nuestro tiempo compartido). Al promover conversaciones significativas, quisimos sentar las bases para temas clave y acciones concretas de cara al futuro.

Llevamos a cabo la conversación de forma virtual, con interpretación simultánea en inglés, portugués y español. Quisimos que todo el mundo pudiera expresarse en una lengua en la que estuvieran cómodos, y que pudieran compartir sus experiencias y saberes con el resto sin que hubiera barreras lingüísticas.

Ha sido una conversación en la que aprendimos muchísimo. Pudimos escuchar las voces provenientes de diferentes disciplinas académicas, el diálogo entre conocimientos —que siempre resulta tan enriquecedor—, y no solo eso, sino también entre muchos países, con muchas lenguas. Un pequeño Babel.

MONICA CECILIA CALDERON

SESIÓN DE CLAUSURA, 13 DE OCTUBRE DE 2021

Además, establecimos prácticas claras en torno a la privacidad para garantizar la seguridad y el bienestar de todos durante el evento. Por ejemplo, los participantes podían tener las cámaras apagadas si así lo preferían, y tenían que pedir consentimiento si querían publicar y citar cualquier declaración de otro participante antes de compartirla en las redes sociales durante y después del evento.

A continuación compartimos cómo se desarrollaron las conversaciones.

Lo que surgió

Perspectivas y provocaciones

La conversación comenzó con una mesa redonda llamada Perspectivas y provocaciones, donde abordamos las cuestiones más urgentes sobre la descolonización de los datos estructurados y cada ponente explicó lo que esto significa desde su punto de vista. La primera parte contó con las aportaciones de Lydia Pintscher, gestora de producto en Wikidata; Matariki Williams, escritora, curadora e

historiadora maorí que actualmente trabaja en el Ministerio de cultura y patrimonio de Aotearoa, Nueva Zelanda; Paz Peña, activista y desarrolladora feminista que se conectó desde Chile y Stacy Allison-Cassin, académica de la Universidad de Toronto, bibliotecaria, entusiasta de Wikidata y miembro de la Nación Métis de Ontario, Canadá. Esta primera mesa redonda sentó las bases para saber qué hay en juego cuando hablamos de datos estructurados.

Empezamos por situar las preguntas y los marcos que emplearíamos a lo largo del evento: qué son los datos estructurados, y qué significa construir sistemas en torno a los mismos, centrados en múltiples epistemologías. Para navegar estos diálogos, hicimos hincapié en el análisis de poder: cuestionamos cómo se manifiesta el control de los recursos —desde el diseño de las plataformas hasta el acceso al contenido en cada una de las lenguas— en las diferentes capas de internet. Preguntas como «¿la historia de quién?», «¿quién hace de narrador y curador?», «¿quién marca los sistemas de categorización?» atravesaron muchas de las intervenciones. Conforme fue

avanzando la mesa redonda, nos ocupamos de la importancia de los datos estructurados y preparamos el terreno para imaginar otras posibilidades.

En la primera parte del evento, los ponentes cuestionaron la supuesta neutralidad de los datos estructurados. Abordamos los datos estructurados como instrumentos de una ideología y no como categorías neutras que clasifican el mundo de una forma determinada. Los participantes navegaron por los matices que surgen cuando cuestionamos la clasificación de los datos y las epistemologías en que se basa dicha clasificación, y cuando priorizamos las necesidades y el derecho de nuestras comunidades a la determinación y a rechazar estas epistemologías. ¿Cómo construimos alternativas basadas en la emancipación y la liberación? Nuestros panelistas lo expresaron de manera muy hermosa: la conversación ha de avanzar más allá de una mera charla, y reconocerse como una cuestión importante que precisa de recursos para que se pueda mejorar y cambiar.

Aquí tenemos una pequeña mirada al debate inicial. Las respuestas se han sintetizado para facilitar su lectura.

¿Qué son los datos estructurados y por qué es importante que hablemos de ello?

Hay muchos sistemas diferentes que interactúan, que estructuran nuestro movimiento por un espacio en diferentes maneras, y a veces esas formas de pensar nos dirigen. [...] Cada vez que nos encontramos con uno de esos sistemas estructurados, también encontramos una pequeña pieza de ideología que conceptualiza el mundo de una manera determinada. Es una interacción entre fragmentos de información o datos y el sistema de información que los envuelve. [...] Por tanto, para mí el primer paso de los datos estructurados está en la pregunta fundamental de cómo estructuramos el mundo que nos rodea, cómo interactuamos con diferentes sistemas de información. Entonces llegamos a un sistema como Wikidata, que me parece maravilloso en múltiples aspectos, como por ejemplo el modo en que puede ser algo colaborativo entre diferentes comunidades, y tener cierta agencia a la hora de cambiar esas estructuras de conocimiento.

Imagen por Whose Knowledge?, CC BY-SA 4.0, via [Wikimedia Commons](#)



— STACY ALLISON-CASIN

¿Qué significa tener múltiples marcos de conocimiento, múltiples marcos epistémicos o epistemologías en el centro de los datos estructurados?

Una de las cosas que he estado pensando es cómo el poder desempeña un papel en los datos estructurados. En mi caso, que soy usuaria más que diseñadora de sistemas de datos estructurados, y especialmente como persona indígena en un país colonizado, esto se manifiesta de manera muy evidente en varios ámbitos en los que he trabajado. [...] No encajamos en el sistema, porque el sistema no ha sido diseñado por nosotres. Esto tiene muchas implicaciones en la manera en que se ha ido compartiendo, recogiendo, recopilando y enriqueciendo el conocimiento a lo largo del tiempo. [...] Cuando hablamos de procesos internos, desde la perspectiva del Ministerio, lo que queremos es generar y estructurar los datos de manera que nuestro trabajo interno sea más eficiente. Pero, ¿qué significa esto para nuestras contrapartes? ¿Esperamos que se ajusten a un patrón que facilita nuestro trabajo, dejando de poner en el centro sus aspiraciones respecto a la gestión de su propio conocimiento? [...] De esta forma, el pensar sobre aquellos enfoques para abordar los datos estructurados con y desde epistemologías múltiples es algo que ya ocurre, si eres una persona indígena es algo que sucede naturalmente.

Imagen por Whose Knowledge?, CC BY-SA 4.0, via [Wikimedia Commons](#)



MATARIKI WILLIAMS —

¿Cómo podemos repensar los datos estructurados, especialmente desde una óptica feminista y anticolonial?

La estructuración de los datos en estos sistemas es una continuación histórica de esa fantasía blanca, eurocéntrica, colonial y patriarcal que asume que los datos son entes que flotan. Flotan por ahí, despojados de sus orígenes, desligados de nuestros cuerpos y de nuestra historia, y por tanto son libres de circular por el mundo como una divisa neutral y universal. [...] También creo que los datos estructurados muchas veces actúan como una fuerza hegemónica, porque literalmente arrastran formas de pensar, como hacen las fuerzas hegemónicas. Esta fuerza hegemónica se presenta como algo inevitable y aspira a ser la única manera de ordenar el mundo. [...] Si queremos abordar los datos estructurados desde una perspectiva feminista, tenemos que luchar por la democratización del debate.

Imagen por Whose Knowledge?, CC BY-SA 4.0, via [Wikimedia Commons](#)



— PAZ PEÑA

What is one thing you would like to see done differently in structured data today, that will help us come to that space of emancipation and liberation?

En Wikidata estamos intentando ofrecer datos estructurados de manera más justa, igualitaria y representativa de las diferentes experiencias y del mundo en el que vivimos. [...] Estamos intentando lograr un equilibrio entre ofrecer unos datos útiles y usables que las personas puedan emplear para construir aplicaciones y, al mismo tiempo, hacer justicia a las personas; ayudarlas a sentirse representadas y vistas en dichos datos, y contribuir a que estos sean una representación de su mundo. Esto hace que los datos de Wikidata a veces sean más complejos. A menudo nos dicen que nuestros datos son difíciles de utilizar por esto mismo. Si pudiera cambiar algo, sería encontrar maneras de salvar esa brecha.

Imagen por Whose Knowledge?, CC BY-SA 4.0, via [Wikimedia Commons](#)



LYDIA PINTSCHER —

Imaginaciones e implementaciones

En esta sesión, les participantes se dividieron en siete grupos, cada uno de ellos integrado por personas con diferentes trayectorias y ámbitos de especialidad. Cada grupo contó con una moderadora y otra persona encargada de escuchar y tomar notas. En esta ocasión, formulamos preguntas enfocadas al futuro: ¿cómo serían unos datos estructurados con diferentes anclajes y marcos epistémicos? Se pidió a les participantes que imaginaran una internet en 2040 que contara con datos estructurados con estas características.

Los grupos detallaron diferentes elementos en relación con este tema: la necesidad de reconocer que las experiencias vividas van mucho más allá de los datos (estructurados) en línea; la urgencia de crear, promover y expandir iniciativas de divulgación y educación sobre los datos y cómo impactan en la vida de las personas, y las posibilidades que se abren al traer al centro del proceso de diseño, propiedad y liderazgo de las plataformas online a las comunidades marginalizadas.

A continuación presentamos una síntesis de estas conversaciones:



Feliz de participar en las sesiones grupales del evento de @WhoseKnowledge Descolonizar los datos estructurados de internet. El resultado más importante aquí es pensar que aspirar a la justicia en los #GrafosDeConocimiento no debería ceñirse solo a los hechos sino también a los modelos de datos y las limitaciones lógicas.

Escuchar y aprender

Tras los debates grupales, tuvimos una sesión llamada Escuchar y aprender. Invitamos a las personas encargadas de la escucha en cada grupo a compartir en la sala plenaria y, de nuevo, contamos con un plantel espectacular: Thomas Hervé Mboa Nkoudou, científico social, bioquímico y maker camerunés; Maya Indira Ganesh, de India, investigadora y divulgadora tecnológica especialmente en el ámbito de la IA; Jani Pereira, bióloga residente en São Paulo, Brasil; Constanza Verón, historiadora y educadora residente en Buenos Aires y que actualmente es parte de Wikimedia Argentina; Asav Bartov, wikimedista, entusiasta de los datos estructurados y bibliotecario digital; Amanda Jurno, catedrática brasileña e investigadora en ciencias de la comunicación, además de integrante de Wiki Movimento Brasil; y Alex Hanna, socióloga e investigadora en el equipo de IA ética de Google.

Este gran equipo explicó lo que les había sorprendido, emocionado o supuesto un reto y cuáles habían sido las ideas más radicalmente creativas e interesantes, así

como sus propuestas para aplicarlas. Las personas encargadas de la escucha y otros participantes ofrecieron estas ideas clave:

ACCESO Y CONTROL DE LOS DATOS

- ▶ Reconocer que descolonizar los datos estructurados y los sistemas construidos sobre ellos es una cuestión crucial y acuciante de derechos humanos, no simplemente una conversación opcional a tener en los márgenes.
- ▶ Admitir que las comunidades marginalizadas —la mayoría minorizada del mundo— tiene poco o nulo acceso o control a los datos que hoy día gobiernan sus vidas.
- ▶ Compartir conocimientos con diferentes comunidades en torno a los datos, las herramientas técnicas y cómo habitar y utilizar los espacios virtuales y las herramientas digitales de internet de manera plena y segura.
- ▶ Tomar conciencia de que una visión de los datos estructurados centrada en la comunidad quizá sea irreconciliable con las agendas de ciertas personas y organizaciones que

se dedican a los datos estructurados, especialmente de quienes los utilizan anteponiendo el lucro al bienestar humano.

AGENCIA/PARTICIPACIÓN

- ▶ No admitir excusas para no comprometerse con la descolonización de los datos estructurados, creando y compartiendo recursos y construyendo gradualmente un ecosistema donde no haya excusas razonables para no reconocer y representar la diversidad de contextos y epistemologías.
- ▶ Admitir que los pueblos indígenas y otras comunidades marginalizadas deberían tener el derecho a negarse a que sus conocimientos figuren en internet o se conviertan en datos, así como la autonomía para decidir qué información comparten con el mundo.

■ DATOS DISTRIBUIDOS

- ▶ Desarrollar maneras de crear conjuntos de datos más pequeños y conectados, que sean gobernados por las propias comunidades marginalizadas en lugar de por empresas que solo tienen fines lucrativos.

■ UN REDISEÑO CENTRADO EN EL PLANETA

- ▶ Prestar atención a los impactos medioambientales de la infraestructura necesaria para procesar cantidades ingentes de datos (por ejemplo, centros de datos) e incorporarlos en los debates.
- ▶ Reestructurar los sistemas preguntándonos continuamente: ¿Para quién son? ¿Qué propósito cumplen?

VE LOS VIDEOS DE LOS PANELES

[ESCUCHAR Y APRENDER](#) ▶

[SESIÓN DE CIERRE](#) ▶

■ PLURALIDAD DE LOS DATOS

- ▶ Rediseñar sistemas y modelarlos en base a diferentes saberes y a las espléndidas complejidades de las comunidades.
- ▶ Reconocer que la mayor parte de los sistemas de conocimiento del mundo no se basan en textos (y pueden incluir otros medios como canciones, relatos y obras de arte de diversa índole), e ir más allá de lo textual para hallar diferentes formas de codificar los datos.
- ▶ Aprender de los ejemplos locales y sus especificidades para crear cambios más amplios, especialmente porque «local» rara vez implica «pequeño», sobre todo en el ámbito del Sur global.
- ▶ Comprometerse a escuchar e incorporar perspectivas diferentes para cambiar la naturaleza de los datos estructurados, sabiendo que la descolonización es un proceso que conlleva tiempo y esfuerzo por parte de numerosas personas y comunidades.

¿Qué sigue?

Para muchos de los participantes, Descolonizar los datos estructurados de internet supuso una oportunidad muy necesaria para detenernos un momento e imaginar posibilidades radicales para los datos estructurados. También fue una ocasión para hacerlo en comunidad, implicando a diferentes regiones del mundo y disciplinas de estudio y trabajo. De cara a seguir adelante, muchos participantes expresaron su interés en organizar más espacios colectivos como este y dar pasos más concretos hacia unas prácticas emancipadoras.

Como organizadoras del evento, reconocemos la gran envergadura de este reto y sabíamos desde un principio que esta conversación habría de ser la primera de muchas. Al entrar en el año 2022, nos interesa crear y convenir más oportunidades para repensar de manera radical y rediseñar los datos estructurados de internet desde una óptica feminista, anticolonial y antirracista. Con esto en mente, vamos a seguir conectados con todos les

participantes de esta primera conversación mediante una lista de correo electrónico. También mantendremos al tanto a nuestras comunidades amplificadas, instándolas a participar e informándolas a través de [nuestro boletín](#).

Agradecimientos

Descolonizar los datos estructurados de internet no habría sido posible sin todas las personas que participaron en la conversación, mil gracias a les ponentes, escuchantes y participantes. También estamos agradecidas con nuestro

fabuloso equipo de intérpretes, ya que sin ellos esta no habría podido ser una conversación plurilingüe. También queremos expresar nuestra gratitud a quienes conspiraron con nosotras para organizar este evento: Amanda Junto y Érica Azzellini de [Wiki Movimento Brasil](#), Sabine Müller y Dominik Scholl de [Wikimedia Deutschland](#). Por último pero no menos importante, querríamos aprovechar esta oportunidad para dar las gracias a las fantásticas personas que nos apoyan, nuestros colaboradores, socios y aliades en sentido amplio; su apoyo hizo posible esta conversación y hará realidad muchas más.

