

# DESCOLONIZANDO LOS DATOS, UN LENGUAJE A LA VEZ

---

Claudia Magallanes Blanco  
Sabelo Mhlambi  
Nanjala Nyabola

Nick Couldry  
Toussaint Nothias  
Kathleen Siminyu

---

2023

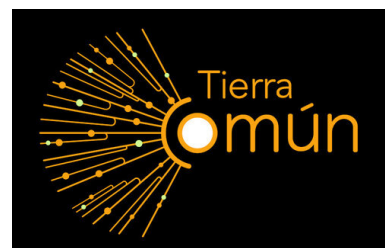


**2023** Claudia Magallanes Blanco, Nick Couldry,  
Sabelo Mhlambi, Toussaint Nothias, Nanjala Nyabola, Kathleen Siminyu.  
Reconocimiento y Compartir Igual.

Traducción al español: **Silvia Coca Córdova**

Traducción al kiswahili: **Alfred Mtawali**

[pacscenter.stanford.edu/digital-civil-society-lab](https://pacscenter.stanford.edu/digital-civil-society-lab)  
[tierracomun.net/](https://tierracomun.net/)



---

# INTRODUCCIÓN

---

En un momento de creciente ansiedad por la ubicuidad de las tecnologías digitales, la idea de que estamos presenciando la era del colonialismo digital y de los datos está ganando fuerza. En los últimos años, personas académicas y activistas han recurrido cada vez más a la noción de colonialidad para entender el impacto social de las tecnologías centradas en los datos, debatir sus daños e imaginar alternativas liberadoras y de empoderamiento.

La cuestión de las lenguas es fundamental en esta área de investigación, diseño y abogacía. En marzo de 2022, el Digital Civil Society Lab (Laboratorio de la Sociedad Civil Digital de la Universidad de Stanford, DCSL por sus siglas en inglés) y la red Tierra Común organizaron una diálogo virtual con ponentes que aportaron diversas perspectivas y puntos de vista sobre el tema. Su trabajo nos invita a reconsiderar el lugar que ocupa el lenguaje en el desarrollo de las tecnologías, pero también en el encuadre público de las mismas. En conjunto, las personas ponentes analizaron por qué es tan importante abordar las brechas lingüísticas en las tecnologías digitales y cómo están arraigadas en un legado colonial. Hablaron de las diversas formas creativas en que las comunidades han respondido a estos desafíos relacionados con la lengua. Y reflexionaron sobre lo que significan en la práctica las perspectivas descoloniales sobre la lengua y los datos.

Este documento es el resultado de esta conversación. En lugar de publicar una grabación del vídeo del evento, hemos preparado una versión editada de la transcripción. Nuestro objetivo es crear cierta fricción generativa en la forma de difundir la información para invitar a las personas a bajar el ritmo y comprometerse de manera más profunda con la conversación. No se escucharán inmediatamente los numerosos acentos en nuestro debate, pero se encontrarán con palabras en diferentes idiomas. Además, publicamos una traducción al español y al kiswahili de la transcripción. La elección de estos dos idiomas refleja el énfasis latinoamericano y africano en las experiencias y los proyectos discutidos por las y los ponentes.

Le agradecemos que dedique su tiempo a este documento y esperamos que le ayude a imaginar visiones compensatorias de lo que puede ser un mundo digital más justo.

Toussaint Nothias

---

## NANJALA NYABOLA:

Muchas gracias a todas las personas por estar aquí. Me encanta y disfruto de la oportunidad de hablar sobre lo que hemos trabajado durante el último año y medio. Gran parte de esto es realmente gracias a la Digital Civil Society Fellowship (beca de la Sociedad Civil Digital de Stanford), la cual hizo posible perseguir algo que había estado en mi mente -y probablemente en la de cualquiera que haya trabajado en tecnología, y ciertamente en derechos digitales, durante los últimos 10-15 años.

La génesis del **Kiswahili Digital Rights Project** (Proyecto de Derechos Digitales en Kiswahili) fue mi propio trabajo en materia de derechos digitales. Como persona multilingüe que creció hablando tres idiomas y estudió un par más en la escuela y en la universidad. Me di cuenta de la brecha existente entre el apoyo lingüístico disponible para las lenguas africanas y el apoyo lingüístico disponible para otras lenguas -las europeas- que tienen muchas menos personas hablantes y mucho menos alcance geográfico. En términos de inversiones para crear un contexto lingüísticamente rico, la brecha era enorme. Y no me refiero realmente a las lagunas técnicas, donde creo que las y los demás ponentes son probablemente más doctas. Me refiero realmente a términos prácticos. ¿Cómo funciona la lengua como una oportunidad, o como un espacio, en el que se invita a las personas a relacionarse con la tecnología en sus propios términos?

Cualquiera que crezca siendo multilingüe te dirá que eres una persona ligeramente

diferente en las distintas lenguas que hablas. Todo el tiempo le digo a la gente que creo que soy más graciosa en kiswahili. Tal vez no sea graciosa en absoluto, pero puedes sacar a relucir diferentes aspectos de tu personalidad, de tus perspectivas sobre la sociedad, de tus perspectivas sobre diferentes cosas, porque el idioma te brinda ciertas palabras. Un idioma pone a disposición ciertas palabras que otro no. Por ejemplo, en mi familia crecimos hablando tres idiomas; utilizamos los tres en una frase, porque la palabra para una acción específica existe en la lengua materna y no existe en kiswahili, y el verbo existe en kiswahili pero no existe en inglés. Es muy típico que las personas multilingües experimenten el mundo de forma diferente debido a esta facilidad lingüística. Y esa fue la génesis personal de este proyecto: darme cuenta de que no podía expresar mi propio interés en los derechos digitales, ni siquiera en kiswahili, que es una lengua enseñada, investigada y bien respaldada. Hay una academia de kiswahili, igual que hay una academia de francés y de alemán. Hay institutos de investigación en múltiples universidades de África Oriental y de Norteamérica y Europa. Sin embargo, existía esta gran brecha para expresar los conceptos fundamentales de los derechos digitales en kiswahili.

Durante los 10-15 años que llevo trabajando en el ámbito de los derechos digitales, escribiendo y organizando en torno a estos diversos temas, siempre nos topamos con esta pared. Por ejemplo, si hacíamos una campaña en torno a Huduma Namba [el programa nacional de identificación digital (ID digital) de Kenia, o "número de servicio",

---

ordenado por el gobierno], que fue el evento detonante, intentábamos explicar a la gente por qué el gobierno no debería poder recopilar sus datos para la ID digital y sin tu consentimiento informado, sin decirte para qué están utilizando los datos y que no los puedan monetizar. Queríamos decir cosas como "protección de datos" y "privacidad de datos", y haríamos esta organización de base y diríamos cosas en kiswahili como "ni muhimu sana kulinda privacy yako" ("es muy importante proteger su privacidad"). Empezaríamos con la palabra inglesa sin su contexto lingüístico, sin ningún tipo de apoyo a su alrededor. Esto ocurre continuamente en varios eventos, pero como dije, el principal evento detonante fue el esfuerzo de 2019 para desplegar este proyecto de ID digital en Kenia, que ocurrió bajo un manto de amenazas e intimidación. A todas y todos los habitantes del país se les dieron 30 días para presentarse en la oficina de su jefe local y registrarse para este ID digital -todas las personas mayores a seis años- si no te registrabas para el ID no podías obtener un pasaporte, un certificado de nacimiento, un certificado de defunción. Básicamente, tu vida como ciudadana o ciudadano dejaría de funcionar.

En el contexto de mi defensa sobre el despliegue del proyecto, seguimos topándonos con esta pared, pues la gran mayoría de las personas kenianas no habla inglés como lengua principal. Aprendemos inglés en la escuela, y es una lengua oficial, pero no es la principal para la comunicación. Y ni siquiera diría que el kiswahili es la lengua principal para la mayoría de las y los kenianos, es la de la interoperabilidad,

que la mayoría de nosotras y nosotros hablamos cuando tenemos que trabajar a través de líneas étnicas, regionales, oficiales y cosas así, así que es la lengua de la interoperabilidad. Pero la mayoría de las personas probablemente se sienten más a gusto malabareando entre las tres. Debido a la imposibilidad de traducirlas correctamente, se suele preferir una sobre otra en función del lugar en el que se encuentre la persona.

Con ese acontecimiento detonante en mente, una idea empezó a tomar forma. Por qué no crear un proyecto para proporcionar las traducciones de algunos de estos conceptos clave: vigilancia, privacidad, protección de datos, consentimiento informado; ninguno de estos términos tenía traducción en kiswahili. Es importante tener en cuenta que el kiswahili es la lengua africana más hablada del mundo. Se habla en al menos 11 países, es oficial en dos, y se enseña en muchos más. Y lo que es más importante, al formar parte de la mayor familia lingüística del continente, también es fundamental para quienes intentan hacer otros trabajos en distintas lenguas. En otras palabras, es mucho más fácil traducir del kiswahili a, por ejemplo, el kikuyu o el banyala, o cualquier otra lengua bantú. Y es más fácil traducir del inglés al kikuyu, o a cualquier otra, porque la estructura gramatical es similar, las filosofías subyacentes de la lengua, las filosofías lingüísticas, son similares. Y luego, hubo un largo año de lucha.

Tuvimos otra catástrofe de derechos digitales con el censo y la idea de que era obligatorio proporcionar a Huduma

---

Namba, que estaba en este proyecto de identificación digital que se estaba desplegando bajo coacción, sin ley de protección de datos, pero también proporcionando coordenadas GPS para todas las personas participantes en el censo, lo cual por supuesto es ilógico pues el censo se supone que funciona para el conteo, no para la identificación. No se supone que deba señalar tu ubicación con coordenadas específicas. Pero no saben realmente cómo empezar, porque esta es otra pieza realmente importante en un contexto de derechos digitales.

Gran parte del discurso está dominado por personas que trabajan en la tecnología o por personas que trabajan en el derecho. Cualquiera de estos espacios responde a las cuestiones del lenguaje. Y aquí no me refiero al lenguaje como una herramienta o un elemento técnico desplegado para construir una mejor tecnología, sino como algo social, como algo construido por la gente a través de las interacciones sociales y las interacciones políticas. Creo que la cuestión del lenguaje era algo que necesitaba la aportación de las y los científicos sociales, de las personas que trabajan en humanidades, pero también de la gente involucrada en las artes creativas. Gente que hace cosas de forma creativa. Por ello llevaba muchos meses queriendo hacer este proyecto -muchos, muchos años en realidad- y no veía realmente la forma de avanzar, pues no existía dentro de estos dos ámbitos en los cuales ya me había desarrollado profesionalmente. Ahí es donde entró la Digital Civil Society Fellowship, porque me proporcionó el financiamiento para poder

realizar este trabajo creativo y lingüístico el cual considero realmente importante para abrir el espacio en la tecnología para pensar en los derechos digitales, no como algo técnico, sino como algo social, como algo que la gente construye a través de su participación y de sus interacciones en el mundo social.

Ha sido, como mencioné, un año y medio. Empezó con talleres con profesoras y profesores de kiswahili de diferentes universidades de Kenia y Tanzania. Esta es la alegría del kiswahili, hay múltiples formas de crear palabras, en algunos casos, se partía de la base de una palabra preexistente. Por ejemplo, "vigilancia", no había ninguna palabra para este término en kiswahili. Antes de hacer este proyecto, utilizábamos la palabra fundacional *kudoea*, cuya traducción más cercana es "codiciar" o mirar algo que alguien tiene y desea. Es como codiciar la información de alguien. Y usamos eso como base, creamos la palabra *udokezi*, que ahora que utilizamos para referirnos a "vigilancia". A veces hay una cosa que hace el kiswahili, se dice *utohozi*, significa básicamente "robar", pero el inglés hace esto todo el tiempo, está bien. Tomas la palabra de otro idioma y la utilizas para romperla estratégicamente en lugares específicos y, en muchos casos, como la gente ya estaba familiarizada con la palabra, *utohozi* tiene más sentido que introducir una palabra completamente nueva, porque, como ya he dicho, el objetivo era crear una lista de palabras que las personas utilizaran. No se trataba de que fuera la palabra más elegante o la más elaborada, sino de hacer cosas que la gente usara.

---

Por ejemplo, *algoridhimu*, que es un "algoritmo"; básicamente la rompimos estratégicamente y cambiamos la pronunciación en lugares específicos para que siguiera sonando familiar, pero fuera una palabra kiswahili.

En algunos casos no teníamos nada de lo que partir y tuvimos que empezar de cero, o alguien más había empezado de cero y básicamente lo usamos para construir la palabra para "programa", *nyavuti*, la cual ya había sido establecida en la gramática en Tanzania, pero no había tenido un uso amplio.

Éstas son las tres categorías básicas que hemos reunido y hemos creado estas increíbles -si me permiten decirlo- tarjetas. En una cara está la palabra en inglés y en la otra la palabra en kiswahili. También las convertimos en naipes. Aquí es donde entra la alegría de trabajar con artistas, pues queríamos hacer algo que la gente usara. Así que cada tarjeta es también un naipe porque hay 54 palabras. Hicimos estas tarjetas y las distribuimos gratuitamente, especialmente entre las personas jóvenes, pues la población africana es generalmente muy joven. Kenia ni siquiera es el país más joven de África. Ni siquiera es el país más joven de África Oriental. El 60% de la población de Kenia tiene menos de 30 años, y todas estas personas han crecido en un mundo en el cual no saben lo que es vivir en un mundo sin Internet. Sin embargo, la educación en materia de derechos digitales no ha estado a la altura. Así que la otra pieza del rompecabezas es, obviamente, la concienciación y la difusión de estas tarjetas lo más ampliamente posible.

El punto que quería dejarles es el siguiente: muchos de los aspectos de los derechos digitales que afectan a la implementación de la tecnología en Kenia no son cuestiones técnicas; son cuestiones políticas y sociales las cuales pueden resolverse mediante innovaciones sociales específicas que apunten a la concienciación, la educación y la inclusión.

Como ejemplo general, ahora mismo la actual administración keniana ha hecho de la "tecnología primero" un pilar central de su política. Si quieres un permiso de conducir, si quieres registrarte para los impuestos o para el pasaporte, tienes que hacerlo en línea. Ninguno de esos sitios web está actualmente traducido al kiswahili. Ni siquiera uno, ¡aunque todos esos sitios web son fundamentales para la ciudadanía! Si tienes que registrarte para alguno de esos servicios, tienes que utilizar esos sitios web.

Y lo más importante, sólo el 12% de las y los kenianos utilizan computadoras personales. Por ende, si una persona tiene que utilizar esos sitios web, lo más frecuente es que deba ir a un cibercafé. Lo cual significa pagar el transporte al cibercafé, pagar a una o un empleado del cibercafé para que introduzca los datos por ti, etc. Básicamente, se trata de un impuesto para las personas que no pueden acceder a estos servicios. Las personas en situación de pobreza están pagando un impuesto para acceder a los servicios del gobierno en un idioma ajeno. Es la lengua de los negocios, pero no de la interoperabilidad. Así que hay todas estas cuestiones silenciosas de derechos que son muy fáciles de articular como cuestiones sociales y no como cuestiones técnicas,

---

pero mientras sigamos centrándonos en la descolonización de la lengua como algo técnico que debe resolverse a nivel técnico, seguiremos dejando atrás a las personas. Y acaba siendo una conversación interna entre gente con poder. Lo que queremos es democratizar el espacio. La descolonización es un proceso activo para recuperar el poder de donde fue robado, de dónde fue interrumpido. Gracias.

---

## KATHLEEN SIMINYU:

Hoy hablaré sobre la construcción de Kiswahili en [Mozilla's Common Voice](#) (La Voz Común de Mozilla). Durante esta presentación les daré una visión general del trabajo de mi equipo y de cómo estamos pensando en los diferentes aspectos de este.

Actualmente soy Machine Learning Fellow (becaria de aprendizaje automático) en la Fundación Mozilla. A grandes rasgos, realizo dos funciones: la primera es organizadora de la comunidad y la segunda es practicante de PLN (Procesamiento de Lenguaje Natural, *NLP* por sus siglas en inglés). Permítanme contarles más sobre el proyecto Common Voice y mi participación en él.

¿Qué es Common Voice? El proyecto se lanzó en junio de 2017. Es el proyecto de Mozilla para crear conjuntos de datos de audio etiquetados, abiertos y disponibles al público, que cualquiera pueda utilizar para practicar aplicaciones de voz. Es parte de los esfuerzos de Mozilla para ayudar a enseñar a las máquinas cómo hablan las personas reales. Es una

apuesta que estamos haciendo por el futuro de la interacción humano-máquina, y es un proyecto que está haciendo que el reconocimiento de voz sea abierto y accesible para todas las personas.

Common Voice trata de recopilar datos de voz de forma segura y ética que puedan utilizarse para entrenar algoritmos de reconocimiento de voz. En lo concerniente al kiswahili, la Fundación Mozilla se ha embarcado con el financiamiento de un proyecto de tres años con el objetivo de crear un corpus de datos lingüísticos en África Oriental. El trabajo principal tiene como objetivo crear habilidades y capacidades, tanto dentro de la Fundación como en la región, para democratizar la tecnología de voz a través del compromiso y la contribución de la comunidad. Este trabajo también apoya el desarrollo de casos prácticos en la agricultura y las finanzas.

Para el kiswahili, en la última versión de los conjuntos de datos, ya tenemos 655 horas de información disponibles. Nuestro objetivo es llegar a tener entre 1000 y 2000 horas en el próximo año. Conscientes de que los sistemas de reconocimiento de voz han sido sesgados con un bajo rendimiento en las mujeres, así como en las poblaciones de mayor edad, estamos trabajando para garantizar que estos grupos demográficos, que probablemente estén subrepresentados en el corpus, se incluyan intencionadamente, especialmente en los puntos de recolección de datos.

Existe una ligera disparidad entre hombres y mujeres, por lo que estamos trabajando para construir una comunidad de voces diversa e inclusiva, accesible para todas y



---

todos, teniendo en cuenta los problemas estructurales que afectan a la desigualdad de género. Esto puede ser la ubicación, la orientación sexual, la educación, la región y el acento. Además, los dos ámbitos en los que hemos decidido centrarnos son la agricultura y las finanzas, que también tienen un gran impacto en las mujeres, pues la mayoría de las personas agricultoras de las comunidades locales son mujeres. Excluir las sería dejar de lado a las personas a las cuales pretendemos impactar con este trabajo.

Contamos con un Gender Action Plan (Plan de Acción de Género) que orienta la forma en cómo prevemos la inclusión en todo el espectro de género. El plan toma en cuenta los retos que dificultan la inclusión, las medidas y acciones que podemos adoptar para abordarlos, también cómo podemos seguir, mediante indicadores y objetivos, si estamos alcanzando la paridad de género. Se trata de un trabajo dirigido por mi colega Rebecca Ryakitimbo, que incorpora un enfoque participativo desde la ideación hasta la recopilación de datos y el desarrollo de casos prácticos, hasta la creación de modelos y el desarrollo de aplicaciones para el usuario final.

Además, hemos trabajado con lingüistas y personas expertas en idiomas, y puede que te preguntes: ¿por qué necesitamos lingüistas? O puede que no, pues Nanjala ha sido un buen punto de partida para mi parte del trabajo. Es para hacer que la tecnología de voz suene más como nosotras y nosotros. Se ha comprobado que los sistemas de reconocimiento de voz están sesgados o se quedan cortos cuando los datos de origen no están equilibrados. El equilibrio implica tener cantidades iguales de

información en las diferentes categorías que están representadas en estos. Así que, de nuevo, la investigación ha demostrado que el reconocimiento del habla es más preciso para los hombres que para las mujeres, y más preciso para las y los individuos menores de 30 años que para las personas mayores de 30 años. Otras características de los datos que pueden presentar un sesgo en las aplicaciones posteriores, si no se equilibran, son los acentos, los dialectos y las variaciones de los idiomas. Para evitar una situación en la que una aplicación desarrollada para una localidad tenga un rendimiento considerablemente peor en otra localidad (que en realidad puede estar sólo a un par de cientos de kilómetros de distancia), necesitamos personas expertas en idiomas que nos ayuden a identificar cuáles son estos diferentes matices y a garantizar que tenemos una representación intencionada en los conjuntos de datos generales del kiswahili.

Lo que hoy conocemos como kiswahili estandarizado tiene una historia controvertida. Les sorprendería saber que entre las personas suahilis existe cierto nivel de disociación. Esto se debe a que no se consultó a las y los lugareños ni se les hizo partícipes del proceso de selección del dialecto a estandarizar.

Lo que hoy conocemos como kiswahili tiene su origen en un dialecto conocido como kiunguja. Y se conocen 23 dialectos del kiswahili, 13 de los cuales son más utilizados que otros. De esta manera, cuando se decidió qué dialectos debían estandarizarse, la decisión fue tomada exclusivamente por misioneros. Se redujo a un debate en el que los diferentes grupos

---

argumentaron a favor de los dialectos que se hablaban en los lugares donde estaban situados. Por lo tanto, se basó en lo que ellos estaban familiarizados y tuvo poco que ver con la gente o las lenguas.

Después, con la decisión de estandarizar el kiunguja, se llegó a la conclusión de que no era lo suficientemente rico desde el punto de vista lingüístico y que necesitaba un vocabulario complementario. Y, una vez más, en esta fase no se contó con la participación de las personas suahilis, lo que dio lugar a un distanciamiento cada vez mayor entre el kiunguja, el dialecto seleccionado para la estandarización, y el kiswahili que surgió tras el proceso de estandarización.

Luego, con el kiswahili estándar llegaron los esfuerzos por impulsar su propagación asegurándose de que fuera la lengua enseñada y utilizada en las escuelas. Esto dio lugar a la inseguridad lingüística y contribuyó a la continua masacre del kiswahili y sus dialectos.

La inseguridad lingüística es la imagen negativa de una persona hablante respecto a su propia variante de habla o lengua. Puede ocurrir si una o un hablante compara sus características fonéticas y sintácticas del habla con las características de lo percibido como correcto, y por desgracia esto es exactamente lo que las y los nativos del swahili sufren tanto en la escuela como en otros entornos formales donde interactúan con las personas que aprendemos y utilizamos el, ahora, kiswahili estándar. Se les corrige repetidamente, haciéndoles creer que lo hablado en casa es incorrecto, pero la verdad es mucho más compleja que eso.

Desgraciadamente, el resultado es un desplazamiento natural hacia lo que se nos hace creer que es correcto, y se aleja de la diversidad presente entre las y los nativos suahilis. En la actualidad, hay dialectos del kiswahili al borde de la extinción y otros que dejan de utilizarse cada día.

Voy a concluir diciendo que de las personas lingüistas y expertas en idiomas aprendemos que construir de forma aislada como personas tecnólogas, como desarrolladoras o incluso como investigadoras de PLN no es lo correcto. Aprendemos que incluso si nuestra intención principal es construir una base de datos en kiswahili en Common Voice y convertirlo en un recurso público disponible con la capacidad de beneficiar a las poblaciones de habla kiswahili, esto no sucederá intrínsecamente si no nos tomamos el tiempo para comprender aspectos de la historia de la lengua y de las propias personas. Aprendemos que si procedemos sin una consideración consciente, corremos el riesgo de alienar a algunas de las poblaciones a las que esta investigación debería beneficiar.

El tiempo compartido con las y los lingüistas consistió en trabajar para identificar los dialectos o variantes dominantes del kiswahili más utilizados en la actualidad. Después, trabajamos con estas personas y les pedimos que hicieran un trabajo de campo útil para elaborar textos que reflejaran las variantes y dialectos que identificamos como predominantes.

En comparación con el trabajo que se está realizando para los conjuntos de datos en kiswahili más amplios, que se basarán en

---

gran medida en el kiswahili estándar, estos subconjuntos serán significativamente más pequeños. Pero nuestra intención es que el texto y el audio recogidos en las respectivas comunidades sean subconjuntos del conjunto. Estos subconjuntos tendrán dos objetivos principales: el primero es ayudarnos a evaluar cuantitativamente el rendimiento de nuestros modelos y aplicaciones posteriores en dialectos y variantes relacionados. Nos gustaría trabajar para conseguir modelos con un rendimiento igual en las distintas variantes y hablantes de dialectos, sin olvidar también sus aspectos de género y edad. Nuestro primer paso será averiguar si efectivamente hay un rendimiento degradado para los distintos grupos. El segundo propósito es que, en caso de que el rendimiento se vea deteriorado para los diferentes grupos demográficos, nos gustaría poner recursos a disposición de las y los desarrolladores para que, dependiendo del contexto local particular para el que estén construyendo aplicaciones, puedan hacer ajustes finos, con el fin de mejorar el rendimiento si es necesario. Muchas gracias.

---

## SABELO MHLAMBI:

Muchas gracias Toussaint, Kathleen y Nanjala por sus comentarios. Me alegro de unirme a todas y todos ustedes aquí.

El proyecto en el que he estado trabajando, especialmente desde que estoy en Stanford, es esta idea de descolonizar el lenguaje desde una perspectiva ética, descolonial y técnica. El lenguaje es el modo en que accedemos al mundo; el

lenguaje contiene visiones del mundo. ¿Qué sucede cuando ya no tenemos el poder de estar a cargo de nuestras propias lenguas?

Me gusta empezar con esta historia de alrededor de 1800. Cuando los misioneros llegaron por primera vez al sur de África, introdujeron un alfabeto basado en la escritura latina. El objetivo principal de este alfabeto era facilitar y abaratar el proselitismo del cristianismo entre las masas africanas. Como probablemente era más barato imprimir más biblias y literatura religiosa que enviar misioneros por mar, tenía sentido. Pero ahora, al estudiar realmente estos alfabetos para, por ejemplo, las lenguas nguni, en Sudáfrica, nos encontramos con que se tarda casi tres veces más en leer una frase o un párrafo en isiZulu que en inglés. ¿Por qué? Porque estos alfabetos no fueron diseñados para satisfacer necesidades. Se diseñaron para difundir un objetivo, que era "¿Cómo convertimos rápidamente a la gente, de la forma más rápida y barata posible?"

Las lenguas bantúes suelen tener una gran morfología, en la que los morfemas (prefijos, sufijos, etc.) se combinan para crear una palabra. Se ha realizado una investigación donde se han seguido los movimientos del globo ocular cuando la gente lee un texto para ver cuánto tiempo se tarda en leer un párrafo en xhosa frente a un párrafo en inglés, y se tarda más del doble en leer el mismo párrafo en xhosa que en inglés debido a la ineficacia de la escritura latina cuando se aplica a las lenguas bantúes. Así que la escritura latina ni siquiera es práctica para comunicarse en las lenguas bantúes. La escritura zulú, por ejemplo, es poco práctica, y la gente suele utilizar las

---

abreviaturas en las redes sociales para escribir rápidamente el idioma...

Al igual que en la época colonial, ahora confiamos en las grandes empresas tecnológicas extranjeras, por ejemplo, para encargarse de dar forma a los mecanismos y las maneras de comunicarnos. Las empresas extranjeras quienes van a crear nuestra tecnología para utilizar nuestras lenguas africanas, y este es el mismo proceso, de nuevo, en el que nos despojamos de este poder, permitiendo que otras y otros creen la forma en cómo vemos y percibimos el mundo. Esto me parece realmente difícil.

He aquí un ejemplo rápido. Tenemos aquí una palabra o frase zulú: *umuntu wabantu*, que significa "un ser humano" o "una persona del pueblo". Pero la traducción de Google dice: "un hombre del pueblo". Espera, espera, ¡nadie ha dicho nada de "un hombre"! Hemos dicho "una persona del pueblo". ¿Por qué se asume automáticamente que es "un hombre del pueblo"? No es eso lo que estamos diciendo. Este es sólo un ejemplo de cómo cuando tenemos sistemas diseñados por otras y otros, pueden moldear nuestra visión del mundo.

Pero, ¿por qué no podemos ser nosotras y nosotros los que creemos nuestras propias tecnologías y tener acceso a estas? Otra amenaza que tenemos con estos sistemas es la vigilancia. Nanjala habló de la vigilancia. Cuanto más información demos sobre nuestras lenguas, mejor podrán vigilarnos estas empresas o hacer un mal uso de nuestra información.

Así que hay otro debate en curso sobre ¿cómo utilizamos nuestra lengua como una

forma de protegernos tácticamente contra la vigilancia o incluso contra las visiones del mundo extranjero que podrían estar en conflicto con nuestra forma de pensar y ver el mundo?. Comencé a pensar mucho en esto, y en los primeros días de mi beca acabé reflexionando más en cómo la mayoría de las y los africanos acceden a Internet o a la tecnología. Es a través de un teléfono móvil, pero eso no es lo suficientemente específico. En realidad es a través de un teclado. Es la puerta de entrada al mundo digital para muchas personas, e incluso en ese ámbito, no tenemos buenos teclados que sean compatibles con las lenguas africanas. Y se trata de herramientas sencillas como la corrección ortográfica. Ni siquiera se puede encontrar esto para el zulú, la segunda lengua más importante de Sudáfrica y la segunda economía más grande del continente africano. Si una lengua así no puede tener ni siquiera recursos sencillos y básicos, ¿cuándo las miles de lenguas que existen en el continente africano tendrán una economía fuerte que los respalde?

Es algo bastante difícil. Si la gente tiene que cambiar lo que es para acceder a la tecnología, para acceder al mundo digital y para aprender otro idioma, entonces ¿cómo puede seguir participando en este mundo digital en el que vivimos?

También empecé a buscar intervenciones técnicas para tratar de abordar lo que se considera una cuestión descolonial. Una de las cosas que hicimos fue crear un teclado, un teclado africano, donde se pueden utilizar GIFs. Tenemos la mayor colección de GIFs africanos y miles de usuarias y usuarios utilizan la aplicación. Tenemos predicciones de palabras, correctores de ortografía, y

---

todo lo que quieras para dar a la gente una forma de comunicarse naturalmente en su propia cultura. Además del teclado tienes una colección de GIFs. Muchos de ellos están hechos a mano, seleccionados y creados para dar voces auténticas a través de las plataformas digitales.

Por ejemplo, si vas a cualquiera de estos sitios web de Giphy y escribes "África", inmediatamente empiezas a ver cosas como una sabana o animales, ¡porque no puedes tener "África" sin ningún animal! Este es el pensamiento general. Una vez más, ¿cómo podemos aprovechar o poseer nuestros lenguajes visuales y cómo queremos expresarnos? Esta es sólo una manera de pensar en hacerlo de manera diferente, por lo que he estado curando esta lista, donde podemos aprovechar eso y crear nuestros propios sistemas.

Me gustaría terminar con esta advertencia, una vez más, si no hacemos caso de las lecciones del pasado, si no aprovechamos nuestras propias lenguas y somos capaces de preservarlas para sostener una parte de nuestra visión del mundo o de nuestros valores, y si no somos capaces de transformar nuestra forma de pensar, incluso al escribir en un papel o en un teclado, lo más probable es que nos quedemos detrás de una brecha digital que se expande constantemente, especialmente a medida que la lengua se convierte cada vez más en la forma de acceder a la tecnología.

Esto es algo que no se ha estudiado ni analizado ampliamente, pero estoy agradecido de poder realizar este trabajo a través de mi beca en Stanford y el DCSL.

---

## CLAUDIA MAGALLANES BLANCO:

Estoy muy contenta de estar aquí y de compartir estas ideas con ustedes. Quiero empezar con algunos conceptos clave como: colonialidad del poder, que se refiere a las interrelaciones de las formas modernas de explotación y dominación; colonialidad del conocimiento, que es el papel de la epistemología y las tareas de producción de conocimiento en la reproducción de los regímenes de pensamiento coloniales; y colonialidad del ser, que es una experiencia vivida de la colonización y su impacto en el lenguaje como el mecanismo donde se inscribe el conocimiento.

La experiencia vivida por las y los sujetos racializados está profundamente marcada por el lenguaje y por cómo éste expresa el mundo como un campo de batalla donde las y los sujetos colonizados son derrotados. El conocimiento que cuenta es el occidental eurocéntrico, cristiano, blanco, masculino y que habla el lenguaje de la metrópoli. El conocimiento no puede expresarse en otro lenguaje que no sea el del poder, porque las y los "otros" no piensan, por lo tanto, no son. ¿Y cómo sabemos que no piensan? Porque no pudieron escribir. No pudieron adentrarse en el mundo de la escritura.

Así pues, la lengua que prevalece está siempre en consonancia con lo que el poder determina como tal, y en conformidad con las políticas lingüísticas dominantes. La lengua es, pues, el soporte de las representaciones y puede convertirse en objeto de luchas por su influencia.

---

Quiero presentarles a Yásnaya Aguilar, una mujer indígena ayuujk y lingüista. Dice: "Nuestras lenguas no mueren solas, nuestras lenguas están siendo asesinadas". Ella considera que la lengua no necesita un espacio para ocurrir, sino que hace que todo ocurra. Hay una interdependencia muy fuerte entre la lengua como territorio cognitivo y la defensa del territorio geográfico, simbólico, material y espiritual. Esto refleja una relación entre la autonomía, el territorio y la lengua. Así, la lingüística permea todo. No es sólo un instrumento de comunicación.

Atacar la lengua es una acción de desposesión del instrumento de pensamiento a través de las violaciones sistémicas de los derechos lingüísticos. Volveré a hablar de Yásnaya más adelante, pero antes quiero introducir el concepto de colonialidad de la lengua, que es un proceso de racialización lingüística que despoja a las poblaciones de su humanidad al atribuirles una inferioridad comunicativa y mental natural. Las lenguas de las personas colonizadas no lo son pues no son capaces de expresar la organización política de un territorio. No son subsidiarias del lenguaje de la "verdad", que eran el latín, el griego o el hebreo. No están escritas en letras y no están constituidas por la relación entre escritura y civilidad. Todos esos imaginarios de las personas colonizadoras sobre las lenguas de las personas colonizadas las convierten en no-humanas porque no tenían esa capacidad lingüística. Las lenguas y las prácticas lingüísticas de las poblaciones colonizadas son formas incapaces de expresar las ideas que las y los colonizadores imaginan como integrales

para ser plenamente humanas. Así pues, en la situación colonial se presupone que las y los colonizados no pueden ser naturalmente interlocutoras e interlocutores.

Ahora quiero pasar de estos marcos de la colonialidad de las lenguas a la presencia de las lenguas indígenas en línea. Nos enfrentamos a una nueva forma de imperialismo, en la que sólo algunas de estas son hegemónicas y existe un dominio pragmático del inglés en el mundo en línea. La exclusión en el contexto capitalista no es sólo no tener acceso a los bienes materiales, también el no tener acceso a la lengua reconocida como de mayor prestigio. Por eso es muy importante la presencia de las lenguas indígenas en línea.

Yásnaya Aguilar dice: "El futuro de las lenguas indígenas también está en el ciberespacio. Y tenemos que empezar a construirlo para que sea un lugar más justo y equilibrado para las personas hablantes de las lenguas del mundo."

Actualmente hay cada vez más páginas en sitios web, aplicaciones y todo tipo de medios digitales que utilizan las lenguas indígenas y las ponen en línea, creando conciencia no sólo para las personas que las hablan, sino para las personas que no las hablan y para que sepan que existen y están presentes.

Quiero poner la atención en el punto de Sabelo sobre los lenguajes comunicativos para la autorrepresentación, que es una noción de Alondra Barba Ramírez. Aunque muchos pueblos indígenas nativos tenían formas de expresión gráfica antes del proceso de colonización, estas formas

---

perdieron la posibilidad de existir frente a la escritura y cuestionamiento del español en las esferas de producción y administración del Estado.

Así, la gramática, las normas, consolidan una lengua alejada del uso, de la vida, para ponerse al servicio del desarrollo comercial y político. Por ejemplo, hay una clara noción de que los textiles contaban historias. Había una gramática en los tejidos y había una lectura y una escritura en este lenguaje que no se reconocía, porque no era el alfabeto escrito.

Alondra Barba Ramírez en realidad habla de cómo estos lenguajes culturales y comunitarios de autorrepresentación utilizan diferentes discursos comunicativos para dar a conocer diferentes dimensiones de la vida de los pueblos indígenas.

¿Cuáles son estos lenguajes comunicativos? Pues la lengua indígena, tanto oral como escrita, pero también las ceremonias, las peregrinaciones, los lugares sagrados, las danzas, la música tradicional, las ofrendas y los objetos ceremoniales.

El territorio, los sonidos de los territorios, los datos se representan en diversos elementos y seres de la naturaleza, la vestimenta, las artesanías, todos estos son lenguajes que están presentes en línea. Están presentes en las redes sociales digitales a través de páginas y perfiles: Facebook, YouTube, Instagram, WhatsApp, TikTok, [y] Twitter.

Los pueblos indígenas de México están utilizando fotografías, vídeos musicales, informes, carteles, audios y memes. Aquí tengo la foto de una persona ganadora

de un concurso de poesía en una lengua indígena y luego puedes encontrar todos los videos de concursos como este. Imágenes así, que son imágenes de una ceremonia sagrada en un sitio sagrado, o la página de Facebook llamada "Hablemos de Hikuri", es un nombre wixárika para el peyote, y es un elemento sagrado de la cultura wixárika. De esta forma están presentes en línea no sólo en la escritura, son todas estas otras formas de lenguaje que provienen de la cosmología o la cosmovisión de los pueblos indígenas.

Quiero terminar con esta noción, "Tequiología", que fue acuñada por Yásnaya Elena Aguilar Gil. Ella considera que la tecnología capitalista se desarrolla sobre la sangre de muchos pueblos indígenas y afrodescendientes, lo cual sabemos con certeza. El saqueo de territorios y cuerpos para el desarrollo tecnológico es un hecho, además de reconocer que las y los individuos racializados cuyos territorios y cuerpos son saqueados no pueden participar plenamente en ese desarrollo tecnológico.

Es importante pensar en la producción de una tecnología que no opera desde un punto de vista extractivista. Teniendo en cuenta el trabajo colaborativo realizado a través del apoyo mutuo, que es una parte muy importante de las culturas indígenas y que representa el trabajo comunal, Yásnaya considera la tequiología como una forma de tecnología colaborativa.

Las comunidades indígenas y el software libre funcionan sobre una base comunitaria de apoyo mutuo, dice Yásnaya. Un ejemplo de este proyecto es Mozilla México y un navegador llamado *Mozilla Native* (Mozilla

---

Nativo), que permite a cualquier persona que hable una lengua indígena participar voluntariamente en la traducción de contenidos digitales. Cada personas que es miembro puede poner las horas que tenga para dedicar al proyecto en línea, traduciendo las diferentes palabras que pueda encontrar en ese servidor. Esta es una propuesta modesta para salvar el mundo: la tequiología, la construcción en común de un internet o tecnología mutua.

Hay muchas personas en el mundo que hablan diferentes lenguas, las cuales no son sólo de forma escrita u oral, sino que tienen muchas, muchas capas. E Internet tendría que reflejar eso. Tenemos que aspirar a la neutralidad de la red, que es complicada en la Internet que tenemos, pero también necesitamos tener más equilibrio y equidad en la Internet en general. Esta es una propuesta concreta que me gustaría poner sobre la mesa.

---

## NICK COULDRY:

Es un honor tener la oportunidad de decir algo después de estas brillantes perspectivas. Permítanme hablar con toda la humildad de alguien cuya lengua materna es el inglés. No puedo separar mi capacidad de hablar con una herramienta de la historia colonial de la dominación del inglés como idioma, así que no lo intentaré.

Permítanme comenzar con un pequeño comentario sobre el papel del lenguaje en la red Tierra Común que copatrocinó este evento, y que co-fundé con Paola Ricaurte y Ulises Ali Mejías. El objetivo de

la red es apoyar las perspectivas sobre el colonialismo de datos desde el Sur Global. Desde el principio fuimos muy conscientes de las cuestiones lingüísticas. La mayoría de las y los miembros de la red son de América Latina, donde su primera lengua es el español o el portugués, no el inglés. Así que trabajamos en tres idiomas. Algunas de nuestras reuniones son en español, algunas de nuestras actividades son en español o portugués, pero, por supuesto, se trata de una batalla continua contra la dominación del inglés como lengua y su profundo legado colonial. Es una batalla en la que es muy difícil avanzar, pero eso es parte de lo que intentamos hacer.

Hemos escuchado algunas perspectivas fascinantes sobre cómo los patrones profundos de exclusión del lenguaje llegan hasta el núcleo del aprendizaje automático. Sabelo habló de lo fácil que es perder el poder sobre la propia lengua, y Claudia habló de la colonialidad del lenguaje. Permítanme añadir un último punto, que quizás subyace en todas estas diferentes y muy importantes perspectivas. Tal y como yo lo veo, las luchas que hemos escuchado descritas para nosotras y nosotros no son únicamente sobre el dominio del inglés, aunque por supuesto eso es fundamental para todo. También son luchas contra las fuerzas que estructuran Internet y la Inteligencia Artificial (IA) de tal manera que amplifican específicamente todas esas desigualdades coloniales e históricas, en torno al habla y muchas otras cosas.

Piensa por un momento en por qué estamos aquí hoy. Es porque subyacente a muchas, muchas perspectivas diferentes y válidas, hay una concentración masiva de poder



---

en el mundo. Un poder sobre el diseño de Internet y esta infraestructura. Un poder sobre el diseño y los recursos del procesamiento de la IA. Un poder sobre los recursos para moderar las lenguas, que existe en algunos lugares, y que se financia mucho más que en otros. Un poder sobre el desarrollo de los modelos de negocio que impulsan la posibilidad misma del Internet. Un poder sobre la economía de las plataformas que nos permiten estar en contacto con las demás personas.

Y como Nanjala planteó tan hermosamente, un poder sobre los lenguajes operativos básicos utilizados para describir lo que estamos haciendo cuando construimos Internet. Todo este trabajo hace algo muy obvio lo cual, creo es importante especificar. Refuerza el poder de una élite principalmente anglófona situada en el Norte Global.

Quiero preguntar si es posible agregar una capa adicional a la lucha que las y los otros oradores han articulado tan bellamente, la lucha más amplia que apenas comienza hoy: construir un espacio configurado de manera diferente, un universo diferente de conexión en línea, que no construya esas desigualdades a diario. Donde el poder de control, supervisión, procesamiento y diseño de cada capa de la estructura de Internet se desplace mucho más cerca de las comunidades locales, a cualquier escala que queramos definir las, comunidades que por lo tanto tendrán la oportunidad de utilizar su propio lenguaje en esos actos de diseño, supervisión, etc. Por el momento estamos muy lejos de esa posibilidad debido a la profunda dislocación del espacio en el que existe el poder sobre

Internet y el espacio en el que vive la mayor parte de la población mundial.

Necesitamos una visión de cómo sería esa lucha más amplia, que conectaría con el resto. Por supuesto, se está llevando a cabo un importante trabajo para descentralizar la infraestructura de Internet, el trabajo del fediverso, mastodonte, y varios otros enfoques, pero creo que sólo enmarcando la lucha más grande de esta manera se hace muy obvio que estamos muy, muy lejos de tener esa lucha común como una prioridad. Tal vez eso es lo que también debemos hacer, aunado a sacar a la luz la enorme violencia registrada en las diferentes luchas de las que han hablado las y los anteriores oradores. Esa es una idea más que quería añadir. Es un honor formar parte de este panel. Gracias.

---

## NANJALA NYABOLA:

Quería aprovechar algo que dijo Kathleen y que también surgió en la conversación sobre nuestro trabajo con el kiswahili. A menudo olvidamos que también hay un elemento de imperialismo en muchas lenguas del Sur Global. El amhárico, por ejemplo, tiene una historia muy singular en Etiopía. Para mucha gente es una lengua imperial, aunque no haya sido imperialista en el sentido de venir de lugares externos como Europa o Italia.

Lo mismo ocurre con el kiswahili, el cual tiene una posición política muy inusual y hay mucha resistencia subconsciente a su uso en muchos contextos en Kenia, por lo tanto tenemos este patois Sheng, que en realidad

---

es más popular. Es la lengua franca, pero es la lengua franca que no está reconocida por el Estado ni por los sistemas oficiales.

Quería poner este punto sobre la mesa pues creo que parte de la energía colonial presente en torno a las prácticas lingüísticas es la incapacidad de ver que las lenguas del Sur Global también pueden ser complicadas. Pueden tener historias políticas y sociales complejas.

Es importante reconocer estas historias sociales en los esfuerzos por hacer que todo el proceso funcione. Hubo muchas conversaciones interesantes en torno a la elección de utilizar el kiunguja frente al kibajuni, frente a cualquiera de los otros dialectos que existen, y hemos recibido muchas críticas de varias personas que tienen la misma perspectiva. Dijeron cosas como: "¿Por qué utilizas esto? ¿Por qué no lo haces en sheng en lugar de hacerlo en suajili? ¿Por qué lo haces en este dialecto en lugar de hacerlo en el otro?". Hay mucho contexto político arraigado. Creo que dar -permítanme hablar de las lenguas africanas, pero sé que también es el caso de América Latina- a las lenguas del Sur Global la oportunidad de ser complicadas, de ser política y socialmente complicadas, y no pretender que el inglés sea la única lengua neutra del mundo, también forma parte del proceso de descolonización.

---

## CLAUDIA MAGALLANES BLANCO:

Estoy completamente de acuerdo con Nanjala. Creo que hay que dejar de romantizar las lenguas indígenas de

América Latina y situarlas en un entorno prístino. También son complejas y políticas, y hay luchas de poder dentro de las diferentes lenguas indígenas y las diferentes formas dialécticas de una lengua en particular, como ya explicaron Nanjala y Kathleen en los casos de África. Lo mismo ocurre en América Latina. Tienes el náhuatl, y tienes varias variantes, o tienes el zapoteco, o el ayuuik, con varias variantes, y hay un protagonismo de ciertos grupos sobre otros grupos basado en las diferencias lingüísticas.

Por eso es importante promover las lenguas indígenas y hacer que reflejen el mundo real en el mundo en línea. Pero este mundo real no es un mundo mágico, encantador y pacífico, -sino que es muy complejo, conflictivo y difícil, y eso lo hace mucho más rico e interesante. Con demasiada frecuencia, en el mundo en línea, cuando se ve una lengua indígena, como en el caso de América Latina, o un dialecto en el caso de los aborígenes de Australia, tiene que ser "perfecto", porque se considera "folclórico". Pero tenemos que poner la dinámica política en el centro para realmente tener la conversación.

---

## TOUSSAINT NOTHIAS:

Gracias, Claudia y todas las demás personas. Voy a plantear una pregunta general a todo el panel, siéntanse libres de responder. Voy a hacerla un poco esquemática para que se pueda debatir. Muchos de los ejemplos que han compartido hoy apoyan la defensa de una mayor inclusión lingüística, como la

---

creación de sistemas de reconocimiento de voz más representativos o teclados más contextualizados. Pero me pregunto si en cierto modo eso no apoya también la idea de la expansión de los datos en todo el mundo. Me pregunto si tienen alguna reacción a esa tensión entre la defensa de una mayor diversidad lingüística y el impulso de la tecnología en todos los espacios de la vida social.

---

## KATHLEEN SIMINYU:

Me gustaría destacar el caso de las personas maoríes. Es algo sobre lo que leí hace un par de semanas, o tal vez meses, las y los maoríes han estado construyendo recursos de audio para su lengua en un intento de preservarla. Así que el objetivo número uno es la preservación de la lengua. Pero cuanto más crece ese repositorio de datos, más interés reciben de las grandes empresas tecnológicas y de otras compañías interesadas en obtener ganancias.

Básicamente, estas empresas están interesadas en explotar ese conjunto de datos y crear tecnologías de reconocimiento del habla. Dado que su intención principal era la preservación de la lengua, creo que en este caso tenían muy claro que la propiedad de los datos era suya, ya que la lengua es un derecho propio.

Creo que, a medida que fomentamos el acceso a más lenguas, o la existencia de más repositorios de datos lingüísticos, también deberíamos educar a las personas hablantes de lenguas sobre la propiedad

y lo que eso significa o puede significar cuando empezamos a crear conjuntos de datos lingüísticos. La propiedad puede significar diferentes cosas. Puede significar una parte de los beneficios si, cuando los datos se utilizan para obtener un beneficio. También puede significar tener derecho a hablar sobre los casos de uso que las comunidades implicadas desean construir frente a las que están en contra pues pueden causar daños o tener algún otro resultado o efecto indeseable para ellas.

Al fin y al cabo, si la lengua que hablas se utiliza para construir una aplicación en las finanzas, significa que tú y esa comunidad lingüística tienen acceso a esos recursos. Pero si se utiliza para construir algo perjudicial, significa que usted y su comunidad lingüística son objeto de ese daño. Hacer que los datos lingüísticos estén más disponibles es una oportunidad para tener una conversación mucho más amplia sobre la gobernanza de los datos y cómo podemos empezar a explicar o ver los datos como algo que pertenece más a las y los propietarios o sujetos de los datos y darles el poder de decidir lo que está bien, lo que no está bien, y darles el derecho a excluirse. Porque no todas las personas tenemos que vivir en mundos hiperconectados y potenciados por la IA en los que, ya sabes, mi asistente de voz lo hace todo.

---

## NICK COULDRY:

Es una pregunta muy difícil la que has planteado pues hay muchas capas de poder sobre Internet, y sobre lo que es Internet, por ende tienes razón. La idea

---

de que todas las personas deberíamos estar conectadas de una manera determinada, y que el aprendizaje automático va a ampliar nuestro conocimiento de todo, es, en sí misma, parte de la ideología de los macrodatos, la cual a su vez es parte de los propios sistemas coloniales que necesitan ser desafiados. Por otro lado, la propia tecnología para el reconocimiento de la voz, la posibilidad de que las máquinas puedan acelerar la traducción de los idiomas para, digamos, permitir la traducción simultánea, es una posibilidad abierta que podríamos imaginar se utilizara en estructuras de poder muy, muy diferentes y eso sería realmente positivo. No creo que haya ninguna razón política para descartar de antemano el uso de determinadas tecnologías. Lo necesario es abrir un mundo donde podamos ver las diferencias que pueden marcar esas tecnologías bajo distribuciones de poder radicalmente diferentes, donde las comunidades tengan realmente la capacidad de influir en sus usos y en su diseño. Eso está muy lejos de lo que tenemos en este momento, pero todos los proyectos de los que hemos oído hablar están tratando de encontrar formas de hacerlo. Por ende lo veo como un enfoque techno-social, con énfasis en lo social.

---

## SABELO MHLAMBI:

Creo que la idea de abrir los datos y las lenguas a todo el mundo en cierto modo es bastante peligrosa, pues también es secreta e íntima. No hay forma de saber cómo se utilizarán nuestras lenguas una vez que estén en manos de las grandes empresas.

Aunque decidan utilizar las lenguas para nuestro "propio bien", para conectarnos mejor, seremos personas excluidas de los beneficios que se derivan de la creación de estos servicios; estamos excluidas de los productos finales de esos servicios. No obtendremos los productos finales en sí. Incluso si nos beneficiamos, no estamos recibiendo la mayor parte del beneficio económico. Por ello creo que una de las perspectivas es averiguar ¿cómo podemos hacernos cargo de nuestros datos, protegerlos y asegurarnos de que nos beneficiamos, como personas africanas o latinoamericanas, de nuestro propio patrimonio cultural?. Nosotras y nosotros creamos estas lenguas, que representan nuestras ideas y pensamientos más íntimos. ¿Cómo nos beneficiamos de ellas, incluso cuando otras empresas dicen que quieren "ayudarnos"? ¿Cómo nos aseguramos de que puedan ayudarnos de forma que conservemos la mayoría y el amplio control del poder económico y la dirección de estas tecnologías?

---

## TOUSSAINT NOTHIAS:

Gracias, Cathleen, Nick y Sabelo. Pasemos a una pregunta del chat, que está relacionada con este último punto: "*¿Qué empresas tecnológicas se han mostrado más abiertas y genuinamente interesadas en colaborar con ustedes para conocer mejor su investigación sobre la colonialidad y la descolonización y pasar a la acción? ¿Qué empresas creen que serán más influyentes en este espacio en el futuro?*"

---

---

## NANJALA NYABOLA:

Quiero asegurarme de responder correctamente a esta pregunta. Obviamente, todo el mundo va a decir que Mozilla, pues creo que es la que más ha hecho (especialmente en el espacio de la lengua kiswahili) para llegar a la gente. Una de las cosas que realicé deliberadamente con el Kiswahili Digital Rights Project, es hacer que todo sea gratuito. En la medida en que se puede dar, se da, y se da a la gente que lo va a usar y se da sin esperar que se use.

Creo que en los países africanos a menudo te encuentras atrapada entre dos cosas: el capital privado que intenta dar la vuelta a lo que sea que estés haciendo para aumentar sus ganancias, y esto habla de la pregunta anterior. Y dos: te quedas atrapada entre las organizaciones sin fines de lucro que te dan dinero con la intención de que conviertas lo que estás haciendo en una especie de campaña de sensibilización. Entre estos dos extremos se acaba con un resultado sesgado.

Por poner un ejemplo sobre el idioma, fui a una escuela católica y durante ese tiempo nuestro profesor se negó a enseñarnos la palabra kiswahili para sexo. No querían enseñarla. Era como si nunca hubiera sucedido, como si no existiera, como si no hubiera que saberla. Vas a la escuela y sales de tu escuela pensando que el kiswahili no tiene una palabra para el sexo. Pero, por supuesto, la tiene. El lenguaje se usa. El lenguaje es contexto. El lenguaje es poder. El lenguaje es intención. De esta forma, quedarse atascada entre los impulsos corporativos y los no lucrativos deja muchas

iniciativas corrompidas. Podemos describir cosas relacionadas con el desarrollo con gran detalle en muchas lenguas africanas por -ejemplo, *maendeleo* ["desarrollo"]- y podemos describirlo en términos comerciales, pero no podemos describirlo realmente en términos sociales, dando a la gente palabras que utilizarán en el día a día.

Al tratar de identificar empresas, instituciones y personas que apoyan este trabajo, es importante pensar fundamentalmente en el contenido de código abierto, la accesibilidad y la forma conveniente de empaquetar las cosas para dar a las personas con el fin de satisfacerlas allí en donde están. Pues en última instancia, a menos que se salga de ese paradigma, se está engendrando una disparidad de poder diferente. Probablemente no sea la disparidad que estabas abordando, pero acaba convirtiéndose en una disparidad de poder diferente, y entonces entran las y los socios. Dejo claro a las y los socios o a las partes interesadas, que nadie gana dinero con esto. Todo esto debe ser lo más justo posible, lo más gratuito posible, y sin ninguna expectativa de uso, sin ninguna expectativa de que luego vayas a rellenar un informe de 50 páginas y me digas cuántas palabras has citado en este caso concreto. La paradoja del desarrollo afectante en muchas iniciativas en esta parte del mundo -estar atrapado entre estos dos extremos- obstaculiza una gran cantidad de trabajo realmente importante porque deja de ser sobre el uso orgánico y la cooperación orgánica y se convierte en representaciones de ciertas ideas de progreso.

---

---

## CLAUDIA MAGALLANES BLANCO:

En México existe la Red de Telefonía Celular Comunitaria, que es una red indígena comunitaria de servicios de telefonía móvil. Y la razón por la que traigo esto a colación en relación con la pregunta es porque uno de los principales debates es, no sólo ofrecer el servicio móvil para llamadas o mensajes de texto, sino sobre la apertura de datos en la red móvil a YouTube y las redes sociales, y un gran aspecto de ese debate fue el lenguaje.

¿Qué va a pasar con nuestra lengua si la abrimos? Las y los adolescentes ya tienen teléfonos móviles y cuando van a los lugares donde tienen conexión con las principales redes corporativas, utilizan sus cuentas de Facebook y de YouTube. Pero cuando vuelven a la comunidad, siguen hablando su lengua indígena. Así que si vamos a tener estas redes sociales dentro de la comunidad, ¿qué pasará con la lengua? Es una pregunta muy válida y todavía no hay una respuesta fácil. Pero al menos esta organización desarrollada por las comunidades indígenas está proporcionando apoyo y construyendo la tecnología y las antenas, etc.

No es sólo una solución técnica, también es una solución social y comunitaria, donde la comunidad dice: "¿Qué queremos? ¿A qué tipo de Internet queremos tener acceso en nuestra comunidad? ¿Cuáles son las cuestiones que nos preocupan en términos de lenguaje, pero también en términos de vigilancia y de control?" Muchas de estas comunidades luchan activamente contra proyectos extractivistas como la minería o

las hidroeléctricas. Entonces, ¿qué pasa con las cuestiones de seguridad si se conectan? Por lo tanto hay un debate y eso provoca que sea una comunidad, o un proyecto comunal: se debe tener esas preguntas. Pueden hacer que la tecnología responda a la primera pregunta de "¿Cómo vamos a conectarnos?" Y luego pueden decir: "Bueno, ahora sabemos cómo conectarnos, pero antes queremos hacerlo de esta manera, y con qué limitaciones, y teniendo en cuenta las cuestiones relacionadas con la vida cotidiana de la comunidad y sus integrantes".

---

## TOUSSAINT NOTHIAS:

Gracias, Claudia. Voy a tratar de incluir aquí dos preguntas de la sección de preguntas y respuestas. La primera pregunta es para Siminyu. *"¿Es fácil o difícil hacer una PLN que sea precisa para varios dialectos del suajili? ¿Qué impacto tendría esto en la sostenibilidad y el consumo de energía, es decir, en el coste de una PLN más precisa frente a la energía consumida para conseguirlo?"* Y la segunda pregunta es: *"Siempre me pregunto, y creo que también está en el libro de Nick Couldry, ¿[cómo] maneja el mundo la cantidad de "contaminación" de datos que queremos poner en línea?"*

---

## KATHLEEN SIMINYU:

Así que la primera pregunta es: ¿es fácil crear un reconocimiento de voz para los distintos dialectos del kiswahili que se utilizan? No estoy muy segura. Aún no hemos

---

llegado al punto de entrenar los modelos de reconocimiento del habla y empezar a evaluar su rendimiento. Lo que sí sé es que somos conscientes de que gran parte del trabajo que estamos haciendo ahora va a repercutir en el rendimiento de los modelos que desarrollemos. De esta forma cosas como si el número de las personas hablantes que contribuyen es lo suficientemente diverso. ¿Solo nos dirigimos a las mujeres? ¿Sólo llegamos a las y los jóvenes? Va a importar si tenemos personas mayores representadas en los conjuntos de datos. Va a importar si tenemos personas con un acento nairobiense frente a un acento que probablemente sea de la costa, o si tienen un acento que habla de la influencia del francés como lengua que también utiliza el o la hablante, o del inglés, etc. Así que, en términos de rendimiento, nos hemos enfrentado a los retos de tener el kiswahili sanifu como lengua dominante. Nos encontramos con muchas personas colaboradoras de voz que nos dicen que no es así cómo se comunican o interactúan en su día a día. Nos hemos dado cuenta de que si queremos construir herramientas con las cuales la gente pueda hablar y utilizar con facilidad, debemos construir herramientas que manejen o generalicen algo más que el kiswahili estandarizado.

Y al final, todo dependerá de la diversidad que seamos capaces de incorporar a los conjuntos de datos. Desgraciadamente, no puedo responder de forma cuantitativa, pero puedo decir que estamos tratando de crear uno lo más diverso posible con los recursos disponibles en este momento.

En cuanto a la eficiencia energética del proceso de formación o creación de modelos

-espero estar entendiendo bien la pregunta-, lamentablemente la tendencia en PLN, al menos en el estado del arte actual, es que todo el mundo está construyendo modelos cada vez más grandes, y quieren hacerlos mucho más multilingües. Por ende están lanzando datos lingüísticos sin siquiera mirar la calidad de los datos, porque se ha demostrado que incluso la información desordenada en cantidades copiosas da resultados significativos. Así que, por desgracia, estoy bastante segura de que el trabajo realizado contribuye mucho al cambio climático, ya que consume mucha energía.

Me gustaría que pensáramos más en diferentes formas de construir la PLN. ¿El estado del arte consiste en introducir más datos porque eso va a aumentar definitivamente la eficacia o podemos pensar en otras formas de hacerlo? Es algo repetido en las conversaciones en torno a este trabajo: si los conjuntos de datos de otras lenguas estrechamente relacionadas con el kiswahili pueden ser beneficiosos. Hemos visto el uso de modelos pre-formados y el uso del aprendizaje por transferencia. ¿Podemos empezar a pensar en tomar datos de kinyarwanda, que también está muy bien representado en el conjunto de datos de Common Voice, y ver si el pre-entrenamiento nos da una ventaja en kiswahili, frente a, por ejemplo, el pre-entrenamiento en inglés a kiswahili, el cual ha demostrado dar grandes resultados? Así que ya veremos.

Intentamos pensar en formas diferentes de hacer las cosas. No debemos seguir con la cultura actual de utilizar modelos más grandes y esperar luego mejores resultados.

---

## TOUSSAINT NOTHIAS:

Al inicio tuvimos una pregunta: *"Me encanta el encuadre de la descolonización como re-intermediación de las relaciones de poder, más que como cuestiones técnicas. Dado que muchos países de bajos ingresos están comprometidos con el crecimiento de los ingresos y el desarrollo económico, ¿hay otras ideas poderosas o personas que hayan encontrado al pensar en la descolonización del capitalismo global?"* ¡Gran pregunta! Estamos aquí para eso.

---

## NANJALA NYABOLA:

Una de las cosas que realmente me fascinan del proceso de descolonización es pensar más allá del dinero, esto se vincula de nuevo con la respuesta anterior. Hasta que no separemos el dinero y los intereses comerciales de lo que intentamos hacer, seguiremos atascadas en este ciclo concreto.

Lo que hace el dinero es crear un incentivo para la competencia y no incentivar la colaboración. Todas las personas intentan ser las primeras, las mejores, las más, y normalmente con la intención de conseguir algo de dinero a cambio. Así que pensar de forma creativa sobre las razones por las cuales estamos haciendo esto, y sacar el dinero y la viabilidad comercial de la ecuación, es algo realmente importante que podemos hacer. Pues ese es realmente uno de los poderosos impulsos provenientes del proceso colonial: hacer que todas estas determinaciones de valor sobre la vida de

las personas y las experiencias cotidianas de las personas, dependan principalmente del dinero.

Uno de mis pensadores políticos, filósofo y profesional keniano favorito, el reverendo Timothy Njoya, dice: *"Kenia fue colonizada para ser un mercado de materias primas y nunca ha salido de ese proceso"*, y por eso nuestra descolonización está estancada. Muchos países africanos se relacionan con el mundo como un lugar de extracción donde las materias primas pueden ser tomadas, empaquetadas, procesadas en otro lugar y, finalmente, nos son vendidas de nuevo.

Creo que nos dirigimos en esa dirección cuando se trata de datos y tecnología, también, donde vemos a los países del Sur Global como lugares principalmente de extracción y no como lugares donde el valor existe por sí mismo, incluso si no puede ser necesariamente cuantificado en términos de dinero. Y esto enlaza con la cuestión que se está debatiendo en el chat sobre la elección de qué dialecto preservar. ¿Se preserva el que tiene viabilidad comercial? ¿Eliges el que tiene más probabilidades de desaparecer en los próximos 50 años? Para mí, lo lógico sería conservar el que tiene más probabilidades de desaparecer. Aunque sólo tenga cuatro hablantes, se debe proteger el que no va a estar aquí y luego trabajar hacia atrás en relación con el que necesita más ayuda. Pero eso no es lo que estamos haciendo. Estamos empezando por el que más se habla y luego, eventualmente, podríamos llegar al que hablan cuatro personas.



---

Realmente se trata del interés comercial: encontrar una manera de eliminar ese dinero como incentivo de la forma en que pensamos en el trabajo que se debe hacer puede ayudarnos a pensar en cuál será la forma del trabajo descolonial.

---

## TOUSSAINT NOTHIAS:

Gracias, Nanjala. Vamos a hacer una última pregunta que abarca algunas de las otras: *"¿Cómo se puede descolonizar la lengua o preservar una visión del mundo sin adoptar la postura de las y los colonialistas de ser una y un árbitro o una y un curador? ¿Cómo saber si lo que se hace es realmente algo deseado por la comunidad? ¿Y qué ocurre cuando la comunidad es grande y contiene puntos de vista contradictorios?"*

---

## SABELO MHLAMBI:

Porque estoy en la posición en la que estoy decidiendo en nombre de algunas personas usuarias, porque estoy creando conocimiento -mis prejuicios están ahí. Una de las formas en las que trato de afrontarlo es asegurando que el trabajo en sí mismo está alineado con una filosofía guía, este caso es el proceso *ubuntu*, la filosofía *ubuntu*. Si es una forma bien aceptada de pensar en el mundo o en cómo deberían ser las cosas, intento alinearme con ella como primer punto de partida y luego también trabajo con otros grupos. Comparto mi investigación con organizaciones del continente para asegurarme de que estoy alineado con el trabajo realizado por otras

y otros. Es una forma de colaborar con las demás personas. Pero también creo que lo más importante es entender la propia filosofía guía y los valores que la gente atesora en esa sociedad, asegurándote de estar alineada o alineado con eso y luego comprobando tus prejuicios en cada colaboración.

---

## NICK COULDRY:

Me gustaría añadir algo a esto, porque creo que esta persona miembro del público ha hecho una pregunta realmente buena. Estamos hablando de cuestiones fundamentales de poder y el poder normalmente se reproduce, así que ¿cómo evitamos eso?

Creo, el principio clave, partiendo de lo que acaba de decir Sabelo, son las decisiones de diseño y qué diseñar y por qué, y cómo supervisarlos, deben tomarse mucho más cerca de las comunidades reales que en las abstracciones que son las grandes corporaciones. Creo que en una pregunta anterior se cuestionaba sobre qué teóricos nos inspiraron -sólo quiero mencionar a uno que ha sido muy influyente en América Latina y que discutimos en un grupo de lectura de Tierra Común hace unos meses, que fue Ivan Illich, un austriaco que fue a México. Él escribía hace 50 años y atacaba principalmente al capitalismo. No enmarcaba explícitamente su crítica en términos de colonialismo, pero su libro *Herramientas para la Convivencialidad* sigue siendo una forma increíble de enmarcar la cuestión sobre cómo el poder y la toma de decisiones deben estar cerca

---

de los grupos sociales de personas reales. En cuanto lo alejas de los grupos de personas, se abstrae, se vuelve violento, divide y polariza.

Y es una forma tan brillante de captar lo que ha ido mal con la versión de Internet que tenemos y de la cual hoy hemos pensado en formas de corregir. Si nos remontamos a cualquier teoría, creo que *Herramientas para la Convivencialidad* todavía tiene mucho que enseñarnos, y está disponible en muchos idiomas, incluyendo una hermosa primera edición en español.

---

## CLAUDIA MAGALLANES BLANCO:

Para aportar otra inspiración teórica, añadiría también la palabra del brasileño Paulo Freire, menciona que no se puede descolonizar algo desde fuera, sólo se puede hacer desde dentro. Así que no es nuestro trabajo el que descoloniza las lenguas, son las personas cuyas lenguas están colonizadas las que descolonizan las propias.

Y no es algo que acaba de empezar. En el caso de América Latina viene ocurriendo desde hace 500 años. Las lenguas fueron colonizadas, pero siguieron en constante resistencia y en constante desafío a la lengua colonizadora. Si no fuera así, ya no tendríamos lenguas indígenas. Sólo tendríamos el español o el portugués. El punto de partida, continuando con Freire, es concienciar y sensibilizar sobre los mecanismos y la estructura de la colonización y sobre cómo opera en la lengua para, a partir de ahí, encontrar la

manera de avanzar realmente hacia la descolonización de la lengua.

Lo que yo diría es que no se puede descolonizar desde fuera. Todos los esfuerzos y todas las cosas que hemos debatido aquí tienen una base comunitaria. Surgen de los pueblos cuyas lenguas se ponen en línea, o en una base de datos, o se traducen a un lenguaje técnico a partir de su lengua viva real. Hay que tenerlo como una experiencia vivida; algo que se vive y se respira cada día. Ahí es donde se encuentra el punto de resistencia y así se avanza hacia la descolonización.

En muchos casos esto ha ocurrido, aunque la gente probablemente no diría realmente: "Oh, estoy descolonizando el lenguaje". Pero en realidad lo están haciendo, y entonces tal vez tengamos un enfoque diferente y digamos "esto es descolonización". Pero, ya sabes, no es "descolonización" sólo cuando decimos que lo es, sino cuándo realmente está ocurriendo.

---

## TOUSSAINT NOTHIAS:

Gracias, Claudia, ¡es una gran manera de terminar! Gracias a todas las personas por unirse a este seminario en línea, y gracias a nuestras y nuestros ponentes por compartir ideas, lecturas, prototipos, herramientas, comunidades, y, sobre todo, por su energía para reimaginar nuestro actual ecosistema digital.

# ACERCA DE LAS PERSONAS PONENTES

---

**Claudia Magallanes Blanco** es profesora del Departamento de Humanidades de la Universidad Iberoamericana de Puebla, México. Es doctora en Humanidades por la University of Western Sydney (Universidad de Occidente de Sydney), Australia. Académica y activista por la justicia social, lleva más de 15 años trabajando con colectivos y organizaciones relacionadas con la comunicación comunitaria e indígena, así como con movimientos sociales contemporáneos en América Latina. Junto con Charlotte Ryan y Alice Mattoni, es co-editora de la serie de libros *Media and Communication Activism: The Empowerment Practices of Social Movements* (Activismo mediático y comunicativo: Las prácticas de empoderamiento de los movimientos sociales) editado por Routledge. Es coautora del artículo *Digital Inclusion Across the Americas and the Caribbean* (Inclusión digital en las Américas y el Caribe) ("Social Inclusion" 2020: vol 8:2) y fue colaboradora en *Capture on Media and Communications* (Captura sobre Medios y Comunicaciones) editada por Nick Couldry y Clemencia Rodríguez en el informe del *International Panel on Social Progress* (Panel Internacional sobre Progreso Social). Sus intereses de investigación y publicación incluyen el conocimiento indígena y la tecnología, los movimientos sociales y las redes socioculturales digitales, la justicia cognitiva y la justicia de datos.

**Nick Couldry** es sociólogo de los medios de comunicación y la cultura. Es profesor de Comunicación Mediática y Teoría Social en la London School of Economics and Political Science (Escuela de Economía y Ciencias Políticas de Londres), y desde 2017 es profesor asociado en el Berkman Klein Center for Internet and Society (Centro Berkman Klein para Internet y Sociedad) de Harvard. En otoño de 2018 también fue profesor visitante en el Massachusetts Institute of Technology MIT (Instituto de Tecnología de Massachusetts). Dirigió conjuntamente, con Clemencia Rodríguez, el capítulo sobre medios y comunicaciones del informe de 22 capítulos de 2018 del *International Panel on Social Progress*. Es autor o editor de quince libros, entre ellos *The Mediated Construction of Reality* (La construcción mediada de la realidad) (con Andreas Hepp, Polity, 2016); *Media, Society, World: Social Theory and Digital Media Practice* (Medios, sociedad, mundo: Teoría social y práctica de los medios digitales) (Polity 2012) y *Why Voice Matters* (Por qué la voz importa) (Sage 2010). Sus últimos libros son *The Costs of Connection* (Los costos de la conexión) (en coautoría con Ulises Mejías), *Media: Why It Matters* (Los medios: Por qué importan) (Polity: 2019) y *Media, Voice, Space and Power* (Medios, voz, espacio y poder) (Routledge 2020).

**Sabelo Mhlambi** es el fundador de *Bhala*, una startup de IA que democratiza los avances de la IA a millones de personas africanas a través del Procesamiento del Lenguaje Natural de las lenguas africanas y los lenguajes visuales africanos. También es el fundador de *Bantocracy*, una organización de interés público que se centra en la ética ubuntu y la tecnología. Ha sido becario del Berkman-Klein Center for Internet & Society (Centro Berkman-Klein para Internet y Sociedad), becario del Stanford's Digital Civil Society Lab, y becario de Technology & Human Rights (Tecnología y Derechos Humanos) 2019-2020 en el Carr Center for Human Rights Policy (Centro Carr para la Política de Derechos Humanos). El trabajo de Mhlambi se encuentra en la intersección de los derechos humanos, la ética, la cultura y la tecnología y hace hincapié en las perspectivas del Sur Global en la política de IA. La investigación del Sr. Mhlambi examina en particular, las implicaciones de los derechos humanos de la tecnología algorítmica en las comunidades marginadas y propone un nuevo marco ético para gobernar la creación y el uso de la Inteligencia Artificial de manera que maximice el progreso y la armonía social.

**Toussaint Nothias** es el Director Asociado de Investigación del Stanford University's Digital Civil Society Lab, investigador del Stanford Center on Philanthropy and Civil Society (PACS) (Centro de Filantropía y Sociedad Civil de Stanford) y profesor afiliado al Centro de Estudios Africanos. Es un especialista en comunicación que trabaja en el ámbito del periodismo, la sociedad civil y las tecnologías digitales. Su investigación explora el polifacético legado de las relaciones de poder coloniales en las representaciones mediáticas contemporáneas, las prácticas periodísticas, la vida asociativa y las redes digitales. Ha publicado artículos en el *Journal of Communication*, *Boston Review*, *International Journal of Communication*, *Public Books* y *Media, Culture, Society*. Su trabajo actual explora la intersección de las tecnologías digitales y la promoción en contextos africanos, con un enfoque particular en uno de los proyectos más notorios y controvertidos para aumentar la conectividad en el Sur Global: El proyecto Free Basics (Básicos Libres) de Facebook. En su papel de director de investigación del Digital Civil Society Lab, el Dr. Nothias dirige la agenda de investigación del Laboratorio y los principales programas académicos, incluyendo el programa postdoctoral, los cursos y los talleres y eventos académicos.

**Nanjala Nyabola** es una escritora e investigadora independiente con sede en Nairobi, Kenia. Su trabajo se centra en la intersección entre tecnología, medios de comunicación y sociedad. Es licenciada en Estudios Africanos y Ciencias Políticas por la Universidad de Birmingham, tiene un máster en Estudios Africanos y otro en Migraciones Forzadas, ambos por la Universidad de Oxford, así como un doctorado en Derecho por la Facultad de Derecho de Harvard. Ha ocupado numerosos puestos de investigadora asociada, entre otros, en el Centre for International Governance Innovation (CIGI) (Centro de Innovación para la Gobernanza Internacional), el Overseas Development Institute (ODI) (Instituto de

Desarrollo en el Extranjero), el Oxford Internet Institute (OII) (Instituto de Internet de Oxford) y otras instituciones, al tiempo que ha trabajado como directora de investigación en varios proyectos sobre derechos humanos en general y derechos digitales en particular en todo el mundo. Ha publicado en varias revistas académicas, como *African Security Review* y *Women's Studies Quarterly*, y ha colaborado en numerosas colecciones editadas. Nanjala también escribe comentarios para publicaciones como The Nation, Al Jazeera, The Boston Review y otras. Es autora de *Digital Democracy, Analogue Politics: How the Internet Era is Transforming Politics in Kenya* (Democracia digital, política analógica: Cómo la era de Internet está transformando la política en Kenia) (Zed Books, 2018) y *Travelling While Black: Essays Inspired by a Life on the Move* (Viajar siendo una persona negra: Ensayos inspirados en una vida en movimiento) (Hurst Books, 2020).

**Kathleen Siminyu** es una investigadora de IA que se ha centrado en el procesamiento del lenguaje natural para las lenguas africanas. Trabaja en la Fundación Mozilla como becaria de aprendizaje automático para apoyar el desarrollo de un conjunto de datos de voz común en kiswahili y construir modelos de transcripción del habla para casos de uso final en los ámbitos agrícola y financiero. En esta función, se esfuerza por garantizar que la diversidad para las personas hablantes de kiswahili, en términos de edad, sexo, acento y variante lingüística/dialecto, se tenga en cuenta en el conjunto de datos y los modelos creados. En su investigación en PLN, Kathleen ha trabajado anteriormente en la transcripción del habla de las lenguas luhya y ha contribuido a la traducción automática de las lenguas kenianas como parte de Masakhane. Antes de unirse a Mozilla, Kathleen fue coordinadora regional de AI4D África, donde trabajó con las comunidades de AA -aprendizaje automático- y IA en África para dirigir varios programas. Tiene una amplia experiencia como organizadora de comunidades, habiendo coorganizado la comunidad de Nairobi Women in Machine Learning and Data Science (Mujeres en Aprendizaje Automático y Ciencia de Datos de Nairobi) durante tres años y sigue organizando como parte de los comités de la Deep Learning Indaba (Aprendizaje profundo Indaba) y la Masakhane Research Foundation (Fundación de Investigación Masakhane).

# ACERCA DEL DIGITAL CIVIL SOCIETY LAB

---

Las tecnologías digitales están transformando a la sociedad civil y la democracia. Nuestra dependencia de los sistemas digitales requiere nuevos conocimientos sobre el funcionamiento de estas tecnologías y la forma en que la sociedad civil puede utilizarlas de forma segura, equitativa e intencionadamente, y en apoyo de la dignidad humana. El Digital Civil Society Lab -o DCSL por sus siglas en inglés- de la Universidad de Stanford tiene como objetivo comprender e informar a la sociedad civil en un mundo digitalmente dependiente. A través de becas, investigación, eventos y docencia, alimentamos una generación emergente de personas académicas, defensoras de la comunidad, tecnólogas y responsables políticos que dan forma al futuro de la sociedad civil digital.

El DCSL es una iniciativa de investigación del Stanford Center on Philanthropy and Civil Society (PACS).

# ACERCA DE TIERRA COMÚN

---

Tierra Común reúne a personas activistas, ciudadanas y académicas que quieren que los datos sean descolonizados. Nuestro enfoque específico es América Latina, pero nuestro horizonte es el Sur Global, y todo aquel que rechace el colonialismo de datos como la última manifestación de la modernidad del deseo de dominación del Norte Global. La imaginación es nuestra mejor herramienta. Imaginemos un futuro en el que el terreno de la vida humana no implique la extracción de datos que nos discriminan y nos separan de nuestras propias vidas.