



Universitat Autònoma de Barcelona

Facultad de Ciencias Políticas y de Sociología

Máster Oficial en Política Social, Trabajo y Bienestar

Curso 2019-2020


La Mala Educación: Narrativas del Banco Mundial desde un enfoque léxico-relacional


Garcia, Gonzalo


Tutor:

Dr. Francesc Miguel Quesada

Bellaterra (Cerdanyola del Vallès), Junio de 2020

ADVERTIMENT. L'accés als continguts d'aquesta tesi queda condicionat a l'acceptació de les condicions d'ús establertes per la següent llicència Creative Commons:  http://cat.creativecommons.org/?page_id=184

ADVERTENCIA. El acceso a los contenidos de esta tesis queda condicionado a la aceptación de las condiciones de uso establecidas por la siguiente licencia Creative Commons:  <http://es.creativecommons.org/blog/licencias/>

WARNING. The access to the contents of this doctoral thesis it is limited to the acceptance of the use conditions set by the following Creative Commons license:  <https://creativecommons.org/licenses/?lang=en>

INDICE

Resumen	3
INTRODUCCIÓN	3
MARCO TEÓRICO	4
1. Banco Mundial: Origen y Políticas orientadas a la Educación	4
2. Aportes Computacionales a las Ciencias Sociales	7
METODOLOGÍA Y DISEÑO DE ANÁLISIS	9
PROCESAMIENTO DE DATOS	11
RESULTADOS	12
1. Bag of Words	12
2. Redes de términos	13
3. Dendogramas	14
4. Coeficiente de Inclusión Mutua por término y período	15
CONCLUSIONES	18
REFERENCIAS	20
ANEXO	24

La Mala Educación: Narrativas del Banco Mundial desde un enfoque léxico-relacional

Garcia, Gonzalo

Resumen

La construcción de narrativas discursivas ha sido uno de los mecanismos utilizados por el Banco Mundial para lograr el apoyo de políticas educativas de sesgo neoliberal. A través de diferentes marcos estratégicos a lo largo del tiempo, la institución ha generado un grupo de recomendaciones a los Estados Nación del Sur Global sobre políticas educativas que, analizados a través de la lente de diferentes autores, a partir de técnicas cualitativas, arrojaban como resultado que el fondo de las políticas siempre ha sido el mismo.

A partir de la incorporación de técnicas computacionales a las posibles herramientas a disponibilidad de los científicos sociales, es posible analizar grandes volúmenes de datos no estructurados y generar estructuras que puedan ser analizadas estadísticamente. Utilizando el blog del Banco Mundial, es posible analizar los cambios en el léxico utilizado por la institución en la construcción de narrativas para el período 2010-2020.

La utilización de un coeficiente de inclusión mutua nos ha permitido comprobar que si bien el léxico utilizado entre los períodos 2010-2014 y 2015-2020 se ha modificado, los términos asociados a palabras clave como “educación”, “público”, “estudiante”, siguen remitiendo a palabras asociadas a aquellas utilizadas por la ideología neoliberal para describir al sistema educativo en términos de eficiencias de mercado.

Palabras clave: Banco Mundial, Narrativa, Léxico, Redes, Coeficiente Inclusión Mutua

INTRODUCCIÓN

Las palabras tienen poder y, tal como dijo Gramsci, la conquista del espacio cultural es una etapa previa para conquistar tanto el espacio político como el económico (Burawoy, 2014). Por tanto, dotarse de un instrumento que permita naturalizar una narrativa como válida es una herramienta que permite transformar la realidad a partir del establecimiento de un discurso hegemónico (Gramsci, 1971). En las narrativas del Banco Mundial respecto a cuáles son las políticas educativas más provechosas para los diferentes países (pero siempre con un foco en los denominados “países en vía de desarrollo”) encontramos esta lucha por imponer un discurso hegemónico y, puntualmente, ese discurso responde a una ideología neoliberal en lo económico.

Los diversos marcos estratégicos presentados por el Banco Mundial respecto a políticas educativas han girado alrededor de las mismas premisas: menos Estado y más involucramiento del sector privado. Sin embargo, las narrativas utilizadas para enmarcar esos discursos se han modificado a lo largo del tiempo. Por tanto, es posible analizar los cambios en las estructuras narrativas utilizadas como vehículo de la ideología dominante.

Al mismo tiempo, los cambios acontecidos en los últimos 20 años en el sector de la tecnología han proporcionado a los científicos sociales nuevas herramientas y nuevas fuentes de datos para analizar las realidades sociales. Dos enfoques pueden ser destacados: el análisis de redes (a través de técnicas de visualización y cálculo) y la digitalización de corpus textuales (como nuevo origen de información). Esto permite ampliar las metodologías de análisis y coordinar esfuerzos entre enfoques cuantitativos y cualitativos dentro de las ciencias sociales. El análisis reticular de coincidencias nos permite transformar los corpus textuales en relaciones entre términos y analizar los patrones que emergen de los grafos resultantes.

El objetivo de este análisis es poder desmontar la narrativa del Banco Mundial y mostrar cómo utiliza ciertas estructuras textuales para mantener un mismo discurso neoliberal a lo largo del tiempo, a pesar de su prédica de ser un banco orientado al conocimiento y que busca mejorar las condiciones de vida de los individuos más necesitados.

Puntualmente, nos interesa poder responder las siguientes dos preguntas que guiarán la investigación:

(P1) ¿Cómo se ha comportado el discurso del Banco Mundial con respecto a temas englobados en la categoría de Educación en los últimos tiempos?

(P2) ¿Los discursos son esencialmente los mismos o hay un cambio en su tónica y sus fundamentos?

MARCO TEÓRICO

1. Banco Mundial: Origen y Políticas orientadas a la Educación

Los orígenes del Banco Mundial (BM de ahora en adelante) se remontan a la finalización de la Segunda Guerra Mundial y los posteriores pactos de Bretton Woods (Mundy & Verger, 2015). A lo largo del tiempo, lo que inició como un organismo dedicado a préstamos para la reconstrucción de Europa, se fue transformando primero en un prestamista de proyectos enfocados en mejorar las condiciones de los individuos pobres en las naciones en vías de desarrollo, para luego, a partir del llamado “Consenso de Washington” de principios de la década de 1990, convertirse en un actor principal dentro del circuito de préstamos dirigidos hacia la educación, sosteniendo los dictados de la ideología neoliberal provenientes del citado consenso (Heyneman, 2003).

Sin embargo, la serie de crisis financieras y económicas que golpearon fuertemente a los mencionados países, provocaron un aumento de las protestas de diferentes grupos civiles respecto a los dictados del BM en la agenda de “liberalización de la educación”. Es así como a partir de la presidencia de James Wolfensohn (presidente del BM desde 1995 a 2005), y sirviéndose del discurso proveniente de los Objetivos de Desarrollo del Milenio (UNGA, 2000), el BM buscó transformarse hacia un “banco del conocimiento”, donde el foco corresponde a la implementación de políticas sociales basadas en evidencia científica a través de consejos técnicos a los países solicitantes de préstamos (Stone & Wright, 2007), en lo que Tarabini & Jacovkis (2012) han denominado el “post-Consenso de Washington”.

A pesar de que a partir de este período el BM promueve políticas de mayor gasto público en el sector de educación, se expanden simultáneamente los requisitos de objetivos numéricos y mensurables para tener un control de la eficiencia de ese gasto, y un mayor control de cómo los países realizan ese gasto en base a estos objetivos. Se pasa de promover resultados cuantitativos a resultados cualitativos (World Bank, 2005). Y de la mano de esa eficiencia es como vuelve a entrar el sector privado en la Educación, ya que se postula que los mecanismos de mercado logran un mejor uso de los recursos que el BM canaliza a través de los diferentes países.

La crisis financiera de 2008 modificó el conjunto de países objetivos a los cuales el BM canalizaba fondos, con un mayor peso de países de medianos ingresos y, con ello, una modificación de las demandas asociadas a esos préstamos. Es así como bajo la presidencia de Jim Yong Kim (2012-2019) se re-convirtió el discurso del BM con un foco en “learning for all” (Verger et al. 2014), bajo la iniciativa SABER (Systems Approach for Better Educational Results), donde pasan a tener preeminencia un grupo de puntos de referencia estándares para todos los países y un grupo de sistemas de control enfocados en la actuación de los profesores, una medición de la descentralización del gasto y el control del mismo y la provisión privada de servicios educativos, todo ello basado en la “supuesta” evidencia científica provista por el mismo BM (Samoff, 2012).

La educación ha sido un foco a donde el BM ha comprometido una parte cada vez mayormente sustancial de su cartera de préstamos. Sin embargo, estos préstamos suelen estar condicionados a que los países receptores traduzcan ese dinero en políticas educativas avaladas por el expertise y los consejos de los investigadores del BM. Sabemos que ese expertise está sesgado hacia un grupo de políticas de corte neoliberal, donde se busca un aumento de la inserción del sector privado en la educación, una menor participación del Estado y una flexibilización tanto de los esquemas de estudio como de la situación laboral del profesorado (Klees, 2017).

Para lograr su cometido, el BM no sólo tiene a su disposición métodos directos como lo pueden ser la presión política ejercida a través de la deuda contraída, sino que también recurre a estrategias soft como lo es la construcción de una narrativa discursiva que le permita influenciar las decisiones de aplicación de políticas educativas (Nordveit, 2012; Steiner-Khamsi, 2010;

Storey, 2000). Esa construcción se da tanto a nivel de reforzar la efectividad de las políticas sugeridas, a través del empleo de un lenguaje científico y bajo la interpretación de que las medidas sugeridas provienen de ensayos estadísticamente comprobados de aplicación universal (Klees, 2012), así como de un lenguaje que invita a la participación y colaboración de todos los actores implicados, si bien luego se comprueba que las colaboraciones no son finalmente tenidas en cuenta (Steiner-Khamsi, 2012; Verger et al., 2019).

Siguiendo los pasos de De Siqueira (2012), podemos observar el camino que ha tomado el discurso del BM a lo largo de los años, a partir de los diferentes marcos educativos estratégicos desarrollados por la institución con las sugerencias de políticas educativas a ser utilizadas por los países objetivo. El análisis permite comprender cómo el BM ha confeccionado un lenguaje de persuasión que no ha modificado sus bases ideológicas en el período de tiempo bajo la lupa, sino que ha recurrido a una técnica que podríamos emparentar con el llamado “rebranding” que suele ser utilizado en las áreas de Marketing como un mecanismo para dotar de un nuevo aspecto hacia el público al producto que, en esencia, sigue siendo el mismo (Merrilees & Miller, 2008). En este caso, el corazón de las políticas de impronta neoliberal siguen estando patentes en las recomendaciones del BM, como las antes mencionadas de flexibilización laboral y menor injerencia pública (Kamat, 2012), pero los términos y palabras asociados a esas políticas se han transformado: se ha pasado de políticas enfocadas en la enseñanza primaria hacia políticas de enseñanza superior, sin cambiar de fondo los términos asociados a esas políticas; se ha pasado de palabra asociadas a cantidades a palabras asociadas a calidades; y, finalmente, se ha pasado a que el término “educación” pase de estar emparentado con palabras como conocimiento hacia palabras como habilidades y destrezas (Nordveit, 2012).

Los cambios políticos en la cúpula del BM, que hemos comentado anteriormente, fueron acompañados por modificaciones en los documentos marco generados para la aplicación y evaluación de políticas públicas. Tal como nos comenta De Siqueira (2012:70), el BM realizó grandes cambios en sus estrategias para la educación en los años 1995, 1999, 2005 y 2010. Cada uno de esos marcos estratégicos fue cambiando el foco de la aplicación final de las políticas educacionales a seguir por los países receptores de los préstamos, pero se puede seguir un hilo conductor en todos ellos: la implementación de políticas de sesgo neoliberal.

En el año 2015 la UNESCO lanzó un marco estratégico de desarrollo sostenible de los países en vías de desarrollo, con foco en la educación a lo largo de la vida de los individuos, inclusiva y de calidad (UNESCO, 2015), a la cual el BM se ha plegado como parte del nuevo giro narrativo en sus políticas educativas. Sin embargo, los cambios asociados a este nuevo framework no son tan diferentes, en esencia, con aquellos ya publicados anteriormente por el BM (Auld et al., 2019).

Si es verdad, tal como dice Parker (1990), que el discurso construye realidad, entonces es necesario ahondar en el análisis del discurso y las estrategias discursivas del BM para construir una realidad que favorece a intereses muy particulares, asociados a los países desarrollados y

ciertos sectores industriales, y no a los países en vías de desarrollo que son los que finalmente terminan aplicando las políticas educativas fomentadas, a pesar de que no existan evidencias de que las mismas puedan tener impactos positivos en la población más vulnerable y más necesitada de ayuda (Vuollo & Herrera, 2016).

2. Aportes Computacionales a las Ciencias Sociales

Tal como afirman Bell et al. (2013:8), la ciencia ficción tiene el potencial de incentivar la innovación tecnológica. Y así como Asimov (2010) imaginó al personaje Hober Mallow utilizando un pequeño dispositivo que le permitía desgranar el discurso del representante de Korell y reducir el mismo a su parte más esencial, de la misma forma podríamos imaginar la creación de una herramienta que sirva para desmontar el andamiaje textual utilizado por el BM (y por otras instituciones similares, como el Fondo Monetario Internacional, o diferentes think tanks) que permita encontrar los verdaderos motivos detrás de las palabras, y que nos permita diferenciar cuáles son los objetivos detrás de las políticas. Trabajos como los de Beltrán Arrieta (2011) apuntan en este sentido: desarrollando herramientas que logren resumir textos de disciplinas específicas a sus componentes más importantes, tal como imaginó Asimov.

Actualmente se constata que las publicaciones orientadas a la Sociología en los principales journals académicos están fuertemente diferenciadas según la metodología de base utilizada por el investigador (Traag & Franssen, 2016), no sólo por las técnicas sino también por las referencias utilizadas, en lo que nos atreveríamos a definir como una “endogamia académica”. Sin embargo, frente al avance de las técnicas computacionales y de procesamiento de datos se abre, hoy en día, la posibilidad de hacer uso de algoritmos y procesos de recolección de datos que no estaban disponibles hace tan sólo veinte años atrás. Es así como podemos alternar el uso de datos de orden cualitativo, con datos de orden cuantitativo, a través de la recolección y análisis de los términos comprendidos en los diferentes documentos que funcionan como portadores de las políticas educativas adecuadas según el BM. El surgimiento de una nueva disciplina que está siendo denominada, en principio, como “computational social science”, donde el paradigma principal es que la sociedad puede ser analizada a partir del procesamiento de la información que genera (Cioffi-Revilla, 2014:2), permitiría acercar técnicas estadísticas avanzadas a los científicos sociales, no como una metodología que se convierta en el nuevo cánón de la disciplina, sino para robustecer los análisis cualitativos tradicionales de la sociología con nueva información relevante y nuevos diseños de hipótesis a contrastar (Mützel, 2015), con el objetivo de tomar las características más fuertes de cada método de análisis y lograr investigaciones más profundas.

Tal como dice Watts (2007), hay que pensar cómo la revolución tecnológica del siglo XXI puede fomentar una revolución en las ciencias sociales también. El rápido aumento de texto disponible de manera online, sumado a la digitalización de documentos, expande el volumen de datos que pueden ser utilizados por las ciencias sociales para trabajar nuevas preguntas, o analizar viejas preguntas desde otro ángulo (Edelmann et al., in press; Hopkins & King, 2007).

Metodologías asistidas por computador, en vez de metodologías de recolección manual de datos, permiten a los investigadores resolver tareas complejas y de carga de trabajo manual pesado en poco tiempo.

Esto no quiere decir que sea una mejor técnica, ya que no debemos olvidar que el aporte de sentido proviene de los seres humanos, no de los algoritmos (Cardie & Wilkerson, 2008). Al mismo tiempo, existen problemas éticos que la combinación de tecnología con otras ciencias (como puede ser la física o la biología) no posee: en el caso de la ciencias sociales, existe el riesgo de invadir la privacidad de los individuos (Lazer et al., 2009) y, además, el peligro que emana de la utilización de estas técnicas para modificar el comportamiento de los individuos o para su control social (Brayne, 2017; Mantelero & Vaciago, 2013; Mantelero, 2014). Por lo tanto, como investigadores debemos ser muy cautelosos en los usos y consecuencias de las teorías que surjan a partir de la combinación social y computacional.

La ventaja del presente trabajo es que la narrativa textual del BM está disponible en línea porque, como hemos comentado, es un vehículo más en la estrategia por imponer un relato hegemónico de cuáles son las políticas educativas correctas para los diferentes países. Por lo tanto, no existen problemas éticos del uso de los mismos. Al mismo tiempo, el volumen de contenido es muy amplio como para ser analizado manualmente por un investigador. Por último, el uso de técnicas computacionales nos permiten, como veremos en el próximo apartado, hacer uso de una metodología mixta, utilizando tanto datos cuantitativos como cualitativos y generando un proceso de análisis que puede ser estandarizado y replicado para diferentes narrativas.

Por eso creemos que el análisis de esta investigación tiene relevancia, no sólo como un paso más en la aproximación de las ciencias de la información con las ciencias sociales sino, además, como el proceso de construcción de una herramienta más con la que cuenten los sociólogos (y otros científicos sociales) para analizar narrativas discursivas de diferentes agentes. Y en el caso del objeto de estudio actual, nos permite obtener las claves del discurso desarrollado por el BM, y poder comparar lo que en apariencia se intenta aplicar y lo que en realidad puede llegar a ser aplicado a partir del corpus textual de recomendaciones a lo largo del tiempo que el BM ha presentado en lo que respecta a políticas educativas y las herramientas acercadas por sus expertos a los países receptores.

En base a lo trabajado con anterioridad, articulando a las preguntas de investigación se encuentran las siguientes hipótesis, asociadas a las mismas:

(P1) ¿Cómo se ha comportado el discurso del Banco Mundial con respecto a temas englobados en la categoría de Educación en los últimos tiempos?

(P2) ¿Los discursos son esencialmente los mismos o hay un cambio en su tónica y sus fundamentos?

(H1) La sustancia del discurso del BM no ha variado, sino que lo que se ha modificado es la estructura que representa ese discurso a través de las relaciones de sus términos.

(H2) El fondo esencial del discurso del BM apunta a políticas educativas de ideología neoliberal, sin importar los cambios en los términos utilizados para describir esas políticas.

El modelo de análisis resultante de las hipótesis puede observarse en la ilustración 3 del anexo.

METODOLOGÍA Y DISEÑO DE ANÁLISIS

Este trabajo de investigación va a estar enmarcado en una metodología mixta, tal como está definido en Johnson et al. (2007), con el objetivo de lograr triangulación y complementariedad entre los métodos utilizados (Greene et al., 1989) y donde el enfoque cualitativo será utilizado como generador de hipótesis a ser corroboradas utilizando un enfoque cuantitativo, a partir de un diseño secuencial del modelo de investigación (Schoonenboom & Johnson, 2017). Esto significa que se utilizarán los análisis de discurso realizados por diferentes investigadores respecto a nuestro objeto de estudio, se extraerán las conclusiones de dichos análisis como hipótesis generales iniciales a la metodología cuantitativa, y se realizarán los tests estadísticos pertinentes de forma tal de dar fuerza de prueba a las conclusiones cualitativas. En la ilustración 2 del anexo, y siguiendo las recomendaciones dadas por Mertens (2011) respecto a cómo deben ser presentados los análisis que contengan una metodología mixta, se puede observar el flujo de trabajo y articulación entre los datos de origen cualitativo y aquellos de origen cuantitativo.

El proceso descrito anteriormente es dado en llamarse “quantitizing” (Sandelowski et al., 2009) y si bien existe debate en si por sí mismo constituye un método mixto (Moran-Ellis et al. 2006), entendemos que el flujo de trabajo seleccionado de tomar como fuente de datos cualitativos secundarios los análisis realizados por otros autores y transformar a datos cuantitativos los corpus textuales de fuentes asociadas a esos datos secundarios nos permite hablar de un método mixto. Adicionalmente, se propone también realizar un análisis inverso y, tal como puede observarse en la porción inferior de la ilustración 1 del anexo, transformar las relaciones cuantitativas entre diferentes palabras que surgirá del análisis de grafos y transformarlo en una narrativa, proceso que se denomina “qualitizing” (Onwuegbuzie & Leech, 2019). Entendemos que a partir de la combinación de ambos procesos, dentro de un flujo secuencial continuado, nos permite hablar de una investigación basada en métodos mixtos (Nzabonimpa, 2018). Este meta-análisis, surgido del proceso de “qualitizing”, nos debería permitir reconstruir una narrativa pasible de ser comparada a las conclusiones de De Siqueira (2012) a partir de los grafos obtenidos por medio de la técnica descrita, si nuestra hipótesis de trabajo es cierta.

En cuanto a la operacionalización de las variables a ser analizadas, debemos entender que en este caso se estará trabajando con un modelo reticular, no uno atributivo (López-Roldán & Fachelli, 2015:37) y, por tanto, la unidad de análisis es la relación (López-Roldán & Lozares, 2012), en este caso entre diferentes palabras. Esto se realizará a través de la transformación

de las entradas del blog del BM en un Corpus vectorial que pueda ser tratado computacionalmente. El mismo genera 1 registro por cada 1 entrada, y luego se generará un objeto matriz (Documento-Término) que permite realizar operaciones estadísticas sobre el Corpus (Silge & Robinson, 2017). Esa matriz puede ser transformada en un grafo relacional con una matriz de adyacencia vinculada al mismo (Jurman et al., 2010).

Para analizar los discursos del BM en cuanto a temática educativa, se encuentra a disposición una base de datos no estructurada de información textual: el blog del Banco Mundial, donde sus expertos generan entradas presentando resultados, generando hipótesis de trabajo, etc. Se cuenta con 428 entradas de blog asociadas a la temática de “Education” (dentro del tópico “Education for Global Development”), con un período de tiempo que abarca desde el 16 de abril de 2010 hasta el 8 de marzo de 2020, como puede observarse en el gráfico 7 del anexo. Tomaremos este corpus textual como parte integral del discurso del Banco Mundial, dado que sus expertos son portavoces del discurso tecnocrático de la institución, a partir del cual se busca imponer al mismo como dominante.

El corpus será extraído mediante el uso de un crawler llamado Octoparse (Octoparse, 2017), que permite navegar y descargar los elementos de una página web con una lógica determinada. En este caso, simplemente se ha descargado el título de la entrada del blog, la fecha en la cual fue posteada la entrada, y el texto completo dentro del cuerpo principal del blog. A partir de las fechas obtenidas, y teniendo en cuenta lo antes analizado respecto al nuevo giro que se plantea en la narrativa del BM a partir del marco estratégico publicado por UNESCO en 2015, decidimos generar dos set de datos, con fecha de corte el 1 de enero de 2015, con el objetivo de poder comparar las narrativas resultantes en ambos períodos y testear la hipótesis planteada.

El análisis de las relaciones entre diferentes palabras dentro del corpus textual va a estar enmarcado en lo que Mercado & Tejero (2018:106) denominan “análisis de coincidencias reticular”, donde el objetivo es descubrir relaciones entre objetos (en nuestro caso palabras), que se suceden de manera conjunta. A partir de esas relaciones se construye lo que De la Rosa Troyano (2005:2) define como “diagramas estructurales”. Esto es, una red de nodos interconectados a partir de la citada medida de coincidencia que nos permitirá explorar visualmente la composición de la narrativa del BM respecto a las políticas educativas fomentadas por el organismo y, al mismo tiempo, construir medidas de relación entre nodos, como el coeficiente de inclusión mutua (De la Rosa Troyano, 2005:8), calculado a partir de la siguiente ecuación:

$$0 \leq S(w_i, w_j) = S(k) = \frac{c_{ij}^2}{c_i c_j} \leq 1 \quad (1)$$

donde w_i , w_j son dos palabras, c_{ij} es el número de documentos en que aparecen conjuntamente las palabras w_i y w_j , y c_i es el número de documentos en que aparece la palabra w_i (mismo procedimiento para c_j).

De esta manera se pueden comparar las distancias entre los períodos temporales mencionados anteriormente y las estructuras narrativas que pueden ser extrapoladas a partir del análisis.

Una vez transformado el corpus en grafos, y a partir de las reflexiones de Nordveit (2012), se seleccionarán un grupo de palabras clave: “education”, “learning”, “public”, “private”. Para comprobar nuestra hipótesis: se debería observar que alrededor de cada palabra analizada las palabras más cercanas a la misma son las mismas para ambos períodos temporales, mientras que las palabras que forman parte de una segunda capa de cercanía (mayor coeficiente de inclusión mutua) sí son distintas en cada período.

PROCESAMIENTO DE DATOS

Como se comentó anteriormente, el corpus textual a ser analizado fue obtenido mediante un proceso de crawling que extrajo el cuerpo de los textos-blog de la página del BM. El resultado final de este proceso es un grupo de archivos de texto plano, uno por cada entrada de blog. Los datos de tipo textual adolecen de problemas específicos, distintos de los datos numéricos utilizados generalmente en trabajos estadísticos (Rajaraman & Ullman, 2011). Es por esta razón que varios autores proponen aplicar, dentro de un esquema de trabajo específico para datos textuales, un conjunto de algoritmos que permiten reducir las dimensiones del corpus (Kadhim et al., 2014; Kannan & Gurusamy, 2014; Vijayarani et al., 2015).

Puntualmente, se han utilizado los siguiente métodos de reducción del corpus: (1) eliminación de valores numéricos, (2) eliminación de signos de puntuación, (3) transformación de toda palabra a minúscula, (4) eliminación de palabras más comunes (“stopwords”), (5) stemming.

Las “stopwords” removidas del corpus textual son aquellas recogidas por Dillon & Modha (2001) e implementadas en la librería “tm” del lenguaje estadístico R (Meyer et al., 2008), con el cual fue realizado el análisis completo. En su mayor parte son palabras orientadas a conjunciones, artículos y pronombres.

El proceso de stemming requiere un poco más de explicación. Se llama “stemming” al proceso mediante el cual una palabra es reducida a su “raíz” mediante la eliminación de los prefijos y sufijos adosados a esa raíz (Jivani, 2011), con el objetivo de reducir la diversidad de términos que tienen la misma connotación, pero difieren en su función dentro de la oración. El proceso de stemming no hace diferencia entre palabras que se escriben igual pero tienen una connotación diferente dentro del mismo texto, ya que el objetivo es la reducción de la complejidad del texto para luego poder ser analizado mediante otras herramientas. Sin embargo, existen métodos de reducción de palabras llamados “lemmatizing” que sí toman en cuenta el sentido contextual de las palabras antes de reducirlas a su raíz. Pero en este caso, la puesta en práctica de tales métodos son muy complejos y no todos los lenguajes de

programación cuentan con librerías que lo permitan. En la práctica, se reconocen los posibles errores resultantes del proceso de stemming y se utilizan diferentes algoritmos según el impacto que tenga los errores en el corpus textual bajo análisis. En nuestro caso, se utiliza el algoritmo desarrollado por Porter (Porter, 1980; Porter, 2001), que tiene una tendencia a arrojar falsos negativos (under-stemming) y, por tanto, puede resultar en que dos palabras que en realidad tienen una misma raíz, sean separadas en dos stem distintos. Tomamos esta decisión debido a que la librería que permite realizar lemmatizing no funciona correctamente con el resto de las librerías de R que utilizan el enfoque “tidy” (Silge & Robinson, 2017) utilizado para realizar el análisis y, dada la situación de falta de contextualización, preferimos ser precavidos y no utilizar un método que aumente el coeficiente de inclusión mutua utilizado para medir la distancia entre términos y que nos guíe a dar por válida una hipótesis por sobreestimación.

Por último, dado que para los algoritmos la dupla “world bank” son dos palabras separadas, pero dentro del contexto de análisis de este trabajo tienen valor como representación única de una institución, fue reemplazado dentro del corpus la dupla “world bank” por la conjunción “world_bank”, de manera de ser interpretada por los algoritmos como una única palabra.

En el esquema disponible en el anexo (véase Circuito de Procesamiento 1), puede observarse el paso a paso de la recogida y procesamiento del corpus textual y los pasos realizados para obtener los gráficos y resultados que serán evaluados en el próximo apartado. Al mismo tiempo, el código utilizado para reproducir este trabajo se encuentra disponible online en la siguiente dirección: <https://doi.org/10.5281/zenodo.3836126>. De esta forma, se dispone de toda la información necesaria que garantiza la reproducibilidad de este análisis.

RESULTADOS

El resultado final de los métodos expuestos anteriormente es un objeto llamado “matriz documento-término” que representa para cada documento analizado los términos que aparecen en el mismo y la cantidad de veces que aparecen en él. En nuestro caso, como hemos dividido los datos a ser analizados en dos conjuntos con un punto de corte el 1 de enero de 2015, tenemos por un lado un primer dataset con 164 documentos y 5662 términos que abarca el período 2010-2014, y un segundo grupo de datos con 264 documentos y 7930 términos, abarcando el período 2015-2020.

1. Bag of Words

Es posible analizar en un primer momento cuáles son los términos que son utilizados más veces en ambos períodos. Los gráficos 3 y 4 del anexo nos muestran esta información en formato de “word cloud”, donde a mayor tamaño del término entonces mayor cantidad de veces utilizado en los documentos. En los mismos podemos apreciar que, tal como se puede esperar, las palabras asociadas a “education”, “school” son las más utilizadas para describir

experiencias y políticas educativas. Sin embargo, cuando empezamos a desgranar términos más abajo en la escala, aparecen algunas diferencias: mientras que para el período 2010-2014 las palabras “private”, “teach”, “help” pertenecen al grupo de palabras más utilizadas, para el siguiente período toman preeminencia las palabras “work”, “time”, “train”. También resulta interesante que en el primer período se utilizara más el término “new policy” mientras que para el segundo período desaparece, y suben en uso los términos “skill” y “quality”.

Esta primer aproximación nos permite entender que si bien existen un grupo de palabras muy utilizados dentro de la narrativa discursiva de los técnicos del BM, hay una diferencia entre los dos momentos temporales bajo análisis, que representan marcos estratégicos distintos: uno redactado por el BM, y luego uno redactado por UNESCO y respaldado por el BM.

2. Redes de términos

También es posible, a partir de la matriz documento-término, y utilizando el coeficiente de inclusión mutua como valor aglutinante, armar un grafo de los términos más utilizados y relacionados entre sí para ambos períodos bajo análisis. En este caso, y dado que la cantidad de términos es muy alta y la matriz resultante tiene como propiedad lo que se denomina “densely sparse”, lo que significa que existe una gran cantidad de celdas iguales a cero en la matriz (Davis & Hu, 2011), se ha decidido reducir la cantidad de términos utilizados para representar el grafo. A través de la función “removeSparseTerms” en R, se eliminaron de la representación gráfica aquellos términos que no estuvieran presentes en el 60% de los documentos de cada período. Asimismo, se grafican términos que tienen un coeficiente de inclusión mutua de al menos 0.1 entre sí. Los grafos resultantes son los gráficos 1 y 2 del anexo.

Tal vez el resultado más interesante de comparar ambos períodos es que el término “education” (y sus variantes) no se encuentra en el primer período pero sí en el segundo. Y en este segundo período, la asociación se da con términos esperados como “policy”, “student”, “world_bank”, pero también con términos como “share” y “percentage”, que indican una asociación a términos que indican valores numéricos. Podemos referenciar estas asociaciones a una cada vez mayor injerencia de resultados cuantificables al momento de evaluar las políticas educativas, tal como es de esperar en un set de políticas de sesgo neoliberal (Holborow, 2007; Holborow, 2012), aún bajo el nuevo marco estratégico presentado en el año 2015.

También nos resulta interesante que los otros términos indicados por Nordveit (2012) como utilizados por el BM para marcar aspectos de las políticas impulsadas (léase “learning”, “private”, “public”) no se encuentran en ninguno de los grafos presentados. Esto nos lleva a reflexionar si ciertos términos son utilizados como referencias a partir de las cuales son ancladas las narrativas del BM, pero sin un uso extensivo de las mismas. En cambio, términos como “support”, “system”, “student”, “outcome” y “research” son utilizados en mayor medida por

los técnicos del BM, y alrededor de estos términos es donde hay mayor densidad de relaciones en los grafos. La conjunción de esos términos remiten un lenguaje científico, tal como indica Klees (2012). Entendemos que el corpus textual utilizado para este análisis tiene un mayor sesgo hacia la utilización de esta clase de lenguaje dado que los mayores aportes al blog del BM son realizados por técnicos. Esto también nos indica de que van a existir diferencias entre el análisis realizado sobre los marcos estratégicos de políticas públicas y los textos analizados aquí. El lenguaje utilizado no es el mismo y, por tanto, es posible que las relaciones indicadas en la hipótesis de trabajo no se cumplan en la misma proporción.

Por último, es interesante notar que el grafo del primer período tiene una mayor densidad que el del segundo período (véase tabla VII del anexo), y que en este último hay un subgrafo que une los términos “government”, “policy”, “key”, “community” y “research” que indica parte de la narrativa esperada por Verger et al. (2019) respecto al uso de una narrativa de colaboración entre los diferentes actores involucrados.

3. Dendogramas

Un método exploratorio adicional para buscar relaciones entre términos es calcular los dendogramas resultantes de las matrices documento-término para ambos períodos. Como se puede ver en los gráficos 5 y 6 del anexo, se puede calcular la distancia entre palabras a partir de la citada matriz. Un dendograma es una estructura con forma de árbol que agrupa en un sentido “bottom-up” los términos a partir de la distancia euclidiana entre los términos del corpus. Adicionalmente, es posible agrupar los términos en clusters tomando como referencia un punto de corte o dictaminando la cantidad de clusters que se quiera formar, con un resultante punto de corte asociado. Como nuestro objetivo con este análisis es simplemente explorar posibles relaciones entre términos, no realizaremos una propuesta de clustering asociado a los dendogramas, sino que referenciaremos algunas relaciones interesantes.

Durante el período 2010-2014 encontramos que los términos “world_bank” y “help” están fuertemente relacionados. Si relajáramos la distancia de agrupamiento, estos dos términos se encuentran luego relacionados con un grupo de términos que hacen referencia a palabras que connotan por un lado mediciones de resultados (“achieve”, “results”, “outcome”) y, por otro lado, a palabras asociadas a actores y valores esperados de esos resultados (“support”, “community”, “high”). También es interesante notar que las palabras consideradas más importantes por la literatura antes mencionada (“learn”, “education”, “student”, “teacher”) aparecen separadas del resto de los términos del dendograma, formando clusters propios.

En el caso del período 2015-2020, el término “world_bank” pasa a estar fuertemente asociado en primera instancia con la palabra “project”, y luego con palabras que referencian a estudios cuantitativos y aplicación de políticas (“effect”, “studies”, “implementation”). Estas asociaciones nos dan a entender que a través de la narrativa, el BM busca posicionarse en el imaginario social como un actor de apoyo a proyectos de políticas y la implementación de las mismas,

siguiendo el espíritu de ser un “learning bank” y alejándose del uso de términos técnicos y hacia el uso de términos más políticos. También es interesante notar que en este período las palabras esenciales (“learn”, “education”, “student”, “teacher”) siguen formando sus propios clusters, pero existe una mayor concentración del uso de estos términos. El período 2015-2020 muestra un uso de lenguaje más reducido y repetitivo, a diferencia del período 2010-2014 donde hay una mayor “sparse density” de términos.

4. Coeficiente de Inclusión Mutua por término y período

Finalmente, tal como se ha planteado en la sección de Metodología, es posible calcular el coeficiente de inclusión mutua (CIM) entre diferentes términos para cada período bajo análisis. Se ha calculado la métrica para los términos propuestos a partir del marco teórico (“education”, “learning”, “public”, “private”) y se han añadido dos términos adicionales (“teach”, “student”) a partir de las relaciones encontradas a partir del análisis exploratorio. Se han seleccionado los 5 primeros términos de mayor CIM como representación de los términos más asociados. Las tablas I a VI del anexo muestran los resultados obtenidos. Como puede observarse a simple vista, los términos asociados en cada período temporal no se repiten y, por tanto, la primer hipótesis planteada en este trabajo no se cumple: el vocabulario de las narrativas de los técnicos del BM se modifica con el cambio de marco estratégico general utilizado por la organización.

En la tabla I podemos observar los términos más asociados con la palabra “education” (bajo el stem “educ”) para los períodos 2010-2014 y 2015-2020. Como primer punto, destacamos que las palabras más asociadas han cambiado de un período a otro; ninguno de los términos se repite. Durante el primer período, el término “education” estuvo asociado a palabras que denotan tanto el financiamiento de la educación, la calidad de la misma, y las consecuencias en desigualdad y oportunidades perdidas que pueden ser consecuencia de la falta de políticas educativas por parte de los diferentes países. En cambio, el giro narrativo durante el siguiente período temporal hace referencia a las características que debería tener el sistema educativo (“inclusive”, “equitable”) y con un foco en los estudios superiores. Podemos concluir que en cuanto a los significados asociados a la palabra Educación, la narrativa del BM ha pasado de una especificidad más ceñida a lo analizado por Kamat (2012), a un cambio asociado a tomar como marco de referencia lo propuesto por UNESCO, donde se dejan de lado los temas de financiamiento de las políticas y se reorientan las mismas a la educación superior, en vez de la primaria, tal como venía siendo propuesto por el BM en sus marcos anteriores al 2015.

Por otro lado, la tabla II muestra los términos asociados a “learning” (por medio del stem “learn”) para ambos períodos. De nuevo, podemos observar que ninguna de las palabras se repite. Durante el primer período el aprendizaje se encuentra asociado a palabras que denotan mediciones y evaluaciones de ese aprendizaje (“measure”, “focus”, “smarter”) junto con la palabra “slum” que hace referencia a los barrios pobres de los países a los cuales el BM suele otorgar préstamos y ayudas para políticas educativas. Mientras que durante el segundo período

vemos un cambio asociado a palabras que denotan las condiciones que imperan en el sistema educativo (“meritocrat”, “crisis”, “scatter”, “decent”) y elementos necesarios para el proceso de aprendizaje (“text”). Podemos evaluar un cambio desde la medición de las pautas de aprendizaje con una mirada más tecnócrata, hacia una versión más política de las políticas educativas, analizando las condiciones que imperan entre los individuos que acceden al sistema educativo y los elementos necesarios para garantizar mejores condiciones de aprendizaje.

Los términos “public” y “private” (a través del stem “privat”) son los que esperaríamos, basados en los análisis de los marcos estratégicos del BM por parte de De Siqueira (2012), que fuesen las palabras donde las asociaciones entre los diferentes períodos fuese menor. La dicotomía público versus privado está en el centro mismo de la concepción ideológica del neoliberalismo en cuanto a cómo deben ser enfocadas las soluciones: si desde la lente pública del Estado o a través de la provisión de servicios privados mediados por el mercado. Sin embargo, tal como hemos observado en los términos anteriores, existen diferencias en los vocablos utilizados para referenciar estas palabras (tablas III y IV del anexo). Por parte de lo público, en el primer período se observa presente la dicotomía público-privado con la fuerte asociación del término “privat” pero también se observan palabras que hacen referencia a la apertura del sistema educativo a los actores privados (“concess”, “provid”) como consecuencia de la falta de recursos (“scarce”); o sea, una narrativa de orden neoliberal. En el segundo período hay una mayor injerencia de términos que hacen referencia a los individuos a los cuales se orienta el sistema educativo (“candid”), a los métodos de evaluación (“examin”), pero también a términos de corte neoliberal. La palabra “shadow” hace referencia al precio sombra de los servicios públicos, propio de la microeconomía mainstream, que junto al vocablo “relief” nos indica que si bien los términos utilizados han cambiado, la base ideológica de las políticas propuestas no es tan diferente. Hacer referencia a precios sombras para poder calcular los costes del sistema público educativo tiene como objetivo analizar cómo debería ser la injerencia del mercado en el mismo.

En cuanto al término “private”, sigue asociado a la dicotomía público-privado durante el primer período analizado, junto con palabras que denotan las necesidades de un marco regulatorio que permita funcionar un mercado dentro del sistema educativo para asignar los recursos (“regulatori”, “provid”). Y si bien durante el período 2015-2020 los términos asociados a lo privado han cambiado, sigue vigente la asociación a las referencias de mercado: “return”, “income”, “contract”. Por tanto, aún en el nuevo marco estratégico educativo para este período, se sigue abogando tanto por la injerencia de los privados dentro de las políticas educativas, como por la lógica de ganancias en lo que hace al sistema educativo, lo cual refuerza en este nuevo período bajo análisis los sesgos neoliberales de las narrativas del BM.

El análisis exploratorio de los apartados anteriores nos permitió añadir los términos “teacher” y “student” como una fuente más para constatar nuestras hipótesis. Y si bien los análisis de De Siqueira (2012) mencionan a los maestros como un foco de las políticas neoliberales, no aparecía como un foco prioritario de su análisis (sino asociado a la injerencia del sector privado

en la educación). En este caso, a partir del corpus textual extraído se puede apreciar en la tabla V del anexo, por primera vez vemos una repetición de términos asociados a la palabra “teacher” para ambos períodos: “system” y “effect”. Y si mientras que para el período 2010-2014 el vocablo se relacionaba con términos que hablaban de la relación de este actor con los estudiantes (“time”, “use”, “student”), para el período 2015-2020 aparecen vocablos que indican las prácticas recomendadas a este actor en su forma de ser partícipe del sistema educativo (“train”, “practice”, “learn”). Por tanto, se pasa de un foco a cómo debe ser la relación estudiante-maestro hacia indicar las prácticas consideradas adecuadas del trabajo de docente.

En cuanto al término “student”, mientras en el primer período se asocia a palabras que hacen referencia a términos que buscan medir las competencias de este actor por medio de valores cuantitativos, no holísticos (“perform”, “score”) y los condicionantes dentro del sistema educativo que rodean al estudiante (“school”, “teacher”, “municipal”), para el siguiente período de análisis se refuerza la asociación a vocablos que remiten a evaluaciones cuantitativas del desempeño (“gpa”), pero al mismo tiempo se empieza a asociar al actor a las condiciones de mercado (“tuition”, “fundrais”), con un tendencia a la mercantilización del sistema educativo y la individualización de ese mismo sistema (alineado con una ideología de mercado). Entonces, podemos observar que en el caso de los estudiantes, el nuevo marco estratégico que guía las políticas educativas del BM es aún más profundo en garantizar la implementación de ideas de sesgo neoliberal al menor nivel posible.

Un resumen de lo comentado en los párrafos anteriores puede observarse en la siguiente tabla, donde también se realiza una comparación con los análisis cualitativos ya revisados en la literatura, que posibilitaron la generación de la hipótesis de este trabajo, para la narrativa comprendida para el período 2010-2014.

Tabla Resumen I. Resumen verbal de la narrativa para el período 2010-2014 y comparación con otros análisis cualitativos

	Narrativa Banco Mundial: 2010-2014		Kamat (2012)
	García (2020)	De Siqueira (2012)	
Education	La narrativa gira en torno al financiamiento y la calidad resultante del sistema educativo, y los impactos en cuanto a la desigualdad de oportunidades en la población.	La Educación es vista como una oportunidad de mejorar y/o conseguir habilidades, lo cual permite aumentar la productividad de la población.	La Educación funciona como una generadora de oportunidades de crecimiento para aquellas poblaciones más vulnerables, a través de un enfoque sistémico.
Public	Dicotomía público-privado. Frente a problemas de financiación, se abre la posibilidad de abrir terreno al sector privado.	Dicotomía público-privado. El sector público ofrece servicios educativos, pero es necesario tener en cuenta la transparencia y la responsabilidad resultante del mismo.	
Private	Dicotomía privado-público. Asignación de recursos a través de una regulación del "mercato educativo".	Visión de la educación "market-driven", lo cual habilita la entrada del sector privado en el mismo. Foco en flexibilidad (propia de mercado), ganancias y transferencia de recursos del sector público al privado.	
Learn	El aprendizaje está asociado a la medición y evaluación de resultados. Impregna una mirada tecnócrata de esa dinámica.	El aprendizaje como un elemento que hace al desarrollo económico de un país. Énfasis en resultados que puedan ser medidos y evaluados.	
Teacher	Medida de relación en maestro-estudiante. Foco en habilidades del maestro, como el entrenamiento y la práctica, y su efecto en el sistema educativo de aprendizaje.	Foco en la experiencia previa de los maestros, y el efecto que tiene sobre el desempeño del sistema educativo.	

Fuente: Elaboración propia

CONCLUSIONES

Como parte de su estrategia para implementar políticas educativas de orden neoliberal, el BM cuenta con varias herramientas. Una de ellas es la construcción de una narrativa que oculte los verdaderos objetivos de esas políticas: darle mayor poder de decisión a los actores privados del mercado a costa de debilitar los alcances de los servicios públicos de educación.

Varios autores han analizado los marcos estratégicos que el BM ha presentado a lo largo de los años, sólo para descubrir que frente a los cambios del vocabulario utilizado para describir las sugerencias de mejores prácticas en el orden de políticas públicas educativas, se mantenía de fondo una ideología neoliberal asociada a esas políticas.

En base a las entradas del blog del BM respecto a temas educativos, donde son partícipes los técnicos y evaluadores de esa misma institución, se ha analizado desde una perspectiva cuantitativa, a partir de las hipótesis generadas por otros investigadores en base a metodologías cualitativas, los cambios en los términos y vocablos utilizados por el BM en dos períodos de tiempo: 2010-2014 y 2015-2020, que coinciden con un cambio en los marcos estratégicos impulsados por la institución.

Se ha podido constatar que los términos asociados a las palabras que los investigadores han denominado como más importantes dentro de la construcción narrativa del BM se han modificado entre ambos períodos de tiempo, no pudiéndose constatar la primer hipótesis explicitada en este trabajo. Sin embargo, la segunda hipótesis sí se ha podido corroborar ya que los nuevos vocablos asociados a los términos “education”, “learning”, “public”, “private”, “teacher”, “student”, siguen enfocados en relaciones que denotan políticas neoliberales donde se utilizan medidas cuantitativas para evaluar el nivel de educación por parte de los estudiantes, mención a modelos de financiación de las políticas educativas cedidas a los mercados en vez del Estado, refuerzo de medidas de valor económico y financiero al momento de hablar del sector público y el privado, valores asociados a la meritocracia y la inclusión pero no a la igualdad y la justicia social, mayor control sobre las tareas de los docentes pero no se adicionan mayores apoyos a estos actores. En resumen, el foco neoliberal sigue presente a pesar del cambio de términos utilizados en la narrativa.

Dado que las hipótesis planteadas en este trabajo están basadas en los análisis de De Siqueira (2012) sobre la lectura de los marcos estratégicos elaborados por el BM hasta el año 2010, y a partir del año 2015 hubo un cambio en la dinámica presentada por las políticas apoyadas por la institución al plegarse al marco estratégico de políticas educativas presentado por UNESCO con foco en el período 2015-2030, creemos que es posible que la imposibilidad de constatar la primer hipótesis se deba a esta razón. A partir del resumen verbal elaborado en base a los términos asociados a cada palabra clave para el período 2015-2020 del corpus textual puede notarse un cambio en el vocabulario utilizado comparado con el primer período, pasando de una narrativa más técnica y densa a una narrativa más política y menos densa (observable en las redes de palabras de los gráficos 1 y 2). Sin embargo, sí se ha constatado que la ideología

neoliberal sigue marcando el conjunto de políticas elaboradas por las instituciones globales. Dado que la extensión del trabajo no lo ha permitido, queda como un próximo paso realizar una reconstrucción narrativa extensa y luego un análisis cualitativo de esa narrativa para poder hablar de la aplicación de un proceso de “qualitizing” en sentido estricto; entendemos que este trabajo ha sido una primera aproximación al proceso completo, y queda como trabajo de ampliación construir la narrativa y analizarla con técnicas cualitativas para cerrar el círculo de conexiones de métodos para una metodología mixta.

Creemos que la metodología utilizada en esta investigación puede permitir sacar a la superficie las intenciones de las políticas impulsadas por esta clase de instituciones cuando, desde una estrategia discursiva, hacen recomendaciones a los diferentes Estados, permitiendo sumar técnicas computacionales a las ciencias sociales como una herramienta a disposición de los investigadores como complemento a los análisis de contenido cualitativo clásicos.

REFERENCIAS

- Asimov, I. (2010). *Trilogía de la Fundación*. Barcelona: Debolsillo.
- Auld, E., Rappleye, J., & Morris, P. (2019). PISA for Development: how the OECD and World Bank shaped education governance post-2015. *Comparative Education*, 55(2), 197-219.
- Bell, F., Fletcher, G., Greenhill, A., Griffiths, M., & McLean, R. (2013). Science fiction prototypes: Visionary technology narratives between futures. *Futures*, 50, 5-14.
- Beltrán Arrieta, R. (2011). *Prototipo de sistema para el resumen automático de textos en castellano en un dominio específico*. (Master's Thesis, Universidad Nacional de Colombia). Disponible en línea: <https://pdfs.semanticscholar.org/dbd0/8483d5b6b7fe1f6c62cf5700c8d231c7f438.pdf> [07/04/2020]
- Brayne, S. (2017). Big data surveillance: The case of policing. *American sociological review*, 82(5), 977-1008.
- Burawoy, M. (2014). La dominación cultural, un encuentro entre Gramsci y Bourdieu. *Gazeta de Antropología*, 30(1), artículo 14.
- Cardie, C., & Wilkerson, J. (2008). Special issue on Text Annotation in Political Science. *Journal of Information Technology and Politics*, 5(1).
- Cioffi-Revilla, C. (2014). *Introduction to computational social science*. London and Heidelberg: Springer.
- Davis, T. A., & Hu, Y. (2011). The University of Florida sparse matrix collection. *ACM Transactions on Mathematical Software (TOMS)*, 38(1), 1-25.
- De la Rosa Troyano, F. F., Gasca, R. M., Abril, L. G., & Morente, F. V. (2005). Análisis de Redes Sociales mediante Diagramas Estratégicos y Diagramas Estructurales. *Redes. Revista hispana para el análisis de redes sociales*, 8(1).
- De Siqueira, A. C. (2012). The 2020 World Bank education strategy: Nothing new, or the same old gospel. In *The World Bank and Education* (pp. 67-81). Rotterdam: Brill Sense.
- Dhillon, I. S. & Modha, D. M. (2001). Concept Decompositions for Large Sparse Text Data using Clustering. *Machine Learning*, 42(1), 143-175.
- Edelmann, A., Wolff, T., Montagne, D. & Christopher A. Bail, C.A. (in press). *Computational Social Science and Sociology. Annual Review of Sociology*, Vol. 46. <https://doi.org/10.1146/annurev-soc-121919-054621>
- Gramsci, A. (1971). *Selections from the Prison Notebooks*. New York: International.
- Greene, J. C., Caracelli, V. J., & Graham, W. F. (1989). Toward a conceptual framework for mixed-method evaluation designs. *Educational evaluation and policy analysis*, 11(3), 255-274.
- Heyneman, S. P. (2003). The history and problems in the making of education policy at the World Bank 1960–2000. *International journal of educational development*, 23(3), 315-337.

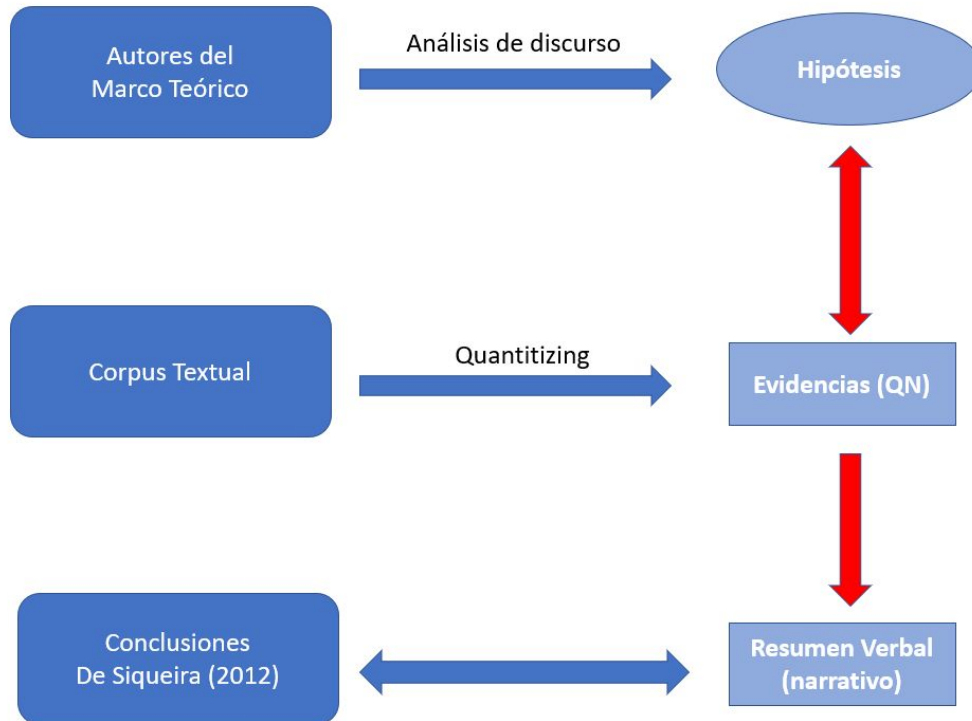
- Holborow, M. (2007). Language, ideology and neoliberalism. *Journal of Language and Politics*, 6(1), 51-73.
- Holborow, M. (2012). Neoliberalism. *The Encyclopedia of Applied Linguistics*, 1-6.
- Hopkins, D., & King, G. (2007). *Extracting systematic social science meaning from text*. Disponible en línea: <http://gking.harvard.edu/files/words.pdf> [04/05/2020].
- Jivani, A. G. (2011). A comparative study of stemming algorithms. *International Journal of Computer Technology and Applications*, 2(6), 1930-1938.
- Johnson, R. B., Onwuegbuzie, A. J., & Turner, L. A. (2007). Toward a definition of mixed methods research. *Journal of mixed methods research*, 1(2), 112-133.
- Jurman, G., Visintainer, R., & Furlanello, C. (2010). *An introduction to spectral distances in networks* (extended version). arXiv preprint arXiv: 1005.0103.
- Kadhim, A. I., Cheah, Y. N., & Ahamed, N. H. (2014, December). Text document preprocessing and dimension reduction techniques for text document clustering. In *2014 4th International Conference on Artificial Intelligence with Applications in Engineering and Technology* (pp. 69-73). IEEE.
- Kamat, S. (2012). The poverty of theory: The World Bank's system approach to education policy. In *The World Bank and Education* (pp. 33-47). Rotterdam: Brill Sense.
- Kannan, S., & Gurusamy, V. (2014). Preprocessing techniques for text mining. *International Journal of Computer Science & Communication Networks*, 5(1), 7-16.
- Klees, S. J. (2012). World Bank and education: Ideological premises and ideological conclusions. In *The World Bank and Education* (pp. 49-65). Rotterdam: Brill Sense.
- Klees, S. J. (2017). Beyond neoliberalism: Reflections on capitalism and education. *Policy Futures in Education*, 18(1), 9-29. <https://doi.org/10.1177/1478210317715814>
- Lazer, D., Pentland, A., Adamic, L., Aral, S., Barabási, A. L., Brewer, D. & Jebara, T. (2009). Computational social science. *Science*, 323(5915), 721-723.
- López-Roldán, P., & Fachelli, S. (2015). *Metodología de la Investigación Social Cuantitativa*. Bellaterra (Cerdanyola del Vallès): Dipòsit Digital de Documents, Universitat Autònoma de Barcelona. 1a. edición. Available at: <https://ddd.uab.cat/record/129382>
- López-Roldán, P., & Lozares, C. (2012). El Atributismo Estructural y el Interaccionismo Estructural en ciencias sociales: ¿concepciones alternativas, antagónicas o complementarias?. *Metodología de Encuestas*, 14, 25-44.
- Mantelero, A. (2014). Social control, transparency, and participation in the big data world. *Journal of Internet Law*, 17(10), 23-29.
- Mantelero, A., & Vaciago, G. (2013). The "Dark Side" of big data: private and public interaction in social surveillance. *Computer law review international*, 14(6), 161-169.

- Mercado, R. M. E., & Tejero, C. (2018). El análisis reticular de coincidencias. *Empiria: Revista de metodología de ciencias sociales*, 39, 103-128.
- Merrilees, B., & Miller, D. (2008). Principles of corporate rebranding. *European Journal of Marketing*, 42(5/6), 537-552.
- Mertens, D. M. (2011). Publishing Mixed Methods Research. *Journal of Mixed Methods Research*, 5(1), 3-6.
- Meyer, D., Hornik, K., & Feinerer, I. (2008). Text mining infrastructure in R. *Journal of statistical software*, 25(5), 1-54.
- Moran-Ellis, J., Alexander, V. D., Cronin, A., Dickinson, M., Fielding, J., Slaney, J., & Thomas, H. (2006). Triangulation and integration: processes, claims and implications. *Qualitative research*, 6(1), 45-59.
- Mundy, K., & Verger, A. (2015). The World Bank and the global governance of education in a changing world order. *International Journal of Educational Development*, 40, 9-18.
- Mützel, S. (2015). Facing big data: Making sociology relevant. *Big Data & Society*, 2(2), 2053951715599179.
- Nordtveit, B. H. (2012). World Bank poetry: How the education strategy 2020 imagines the world. In *The World Bank and Education* (pp. 21-32). Rotterdam: Brill Sense.
- Nzabonimpa, J. P. (2018). Quantitizing and qualitzing (im-) possibilities in mixed methods research. *Methodological Innovations*, 11(2). <https://doi.org/10.1177/2059799118789021>
- Octoparse (2017). *About Octoparse*. Disponible en línea: <https://www.octoparse.com> [08/04/2020]
- Onwuegbuzie, A. J., & Leech, N. L. (2019). On Qualitizing. *International Journal of Multiple Research Approaches*, 11(2).
- Parker, I. (1990). Real things: Discourse, context and practice. *Philosophical Psychology*, 3(2-3), 227-233.
- Porter, M. F. (1980). An algorithm for suffix stripping. *Program*, 14(3), 130-137
- Porter, M. F. (2001). *Snowball: A language for stemming algorithms*. Disponible en línea: <https://pdfs.semanticscholar.org/0d8f/907bb0180912d1e1df279739e45dff6853ee.pdf> [04/05/2020]
- Rajaraman, A., & Ullman, J. D. (2011). *Mining of massive datasets*. Cambridge: University Press.
- Samoff, J. (2012). "Research Shows that...": Creating the Knowledge Environment for Learning for All. In *The World Bank and Education* (pp. 143-157). Rotterdam: Brill Sense.
- Sandelowski, M., Voils, C. I., & Knafl, G. (2009). On quantitizing. *Journal of mixed methods research*, 3(3), 208-222.
- Schoonenboom, J., & Johnson, R. B. (2017). How to construct a mixed methods research design. *KZfSS Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 69(2), 107-131.

- Silge, J., & Robinson, D. (2017). *Text mining with R: A tidy approach*. O'Reilly Media, Inc. Disponible en línea: <https://www.tidytextmining.com/> [04/05/2020]
- Steiner-Khamsi, G. (2010). The politics and economics of comparison. *Comparative Education Review*, 54(3), 323-342.
- Steiner-Khamsi, G. (2012). For All by All?: The World Bank's Global Framework for Education. In *The World Bank and Education* (pp. 1-20). Rotterdam: Brill Sense.
- Stone, D. & C. Wright (2007). *The World Bank and Governance: A Decade of Reform and Reaction*. London: Routledge
- Storey, A. (2000). The World Bank, neo-liberalism, and power: discourse analysis and implications for campaigners. *Development in Practice*, 10(3-4), 361-370.
- Traag, V., & Franssen, T. (2016). *Revealing the quantitative-qualitative divide in sociology using bibliometric visualization*. CWTS blog.
- UNESCO (2015). *2030: Incheon Declaration and Framework for Action for the Implementation of Sustainable Development Goal 4. Ensure inclusive and equitable quality education and promote lifelong learning opportunities for all*.
- United Nations General Assembly (2000). *United Nations Millennium Declaration*. UN Resolution A/RES/55/3. New York: United Nations
- Verger, A., Edwards Jr, D. B., & Altinyelken, H. K. (2014). Learning from all? The World Bank, aid agencies and the construction of hegemony in education for development. *Comparative Education*, 50(4), 381-399.
- Verger, A., Fontdevila, C., Rogan, R., & Gurney, T. (2019). Manufacturing an illusory consensus? A bibliometric analysis of the international debate on education privatisation. *International Journal of Educational Development*, 64, 81-95.
- Vijayarani, S., Ilamathi, M. J., & Nithya, M. (2015). Preprocessing techniques for text mining-an overview. *International Journal of Computer Science & Communication Networks*, 5(1), 7-16.
- Vuollo, E. & Herrera, A. P. (2016). *The Development Imaginary: Analyzing development discourse in the World Bank Education Strategy 2020 through contemporary theories of development and education*. (Master's Thesis, University of Oulu). Disponible en línea: <http://urn.fi/URN:NBN:fi:oulu-201606042345> [07/04/2020]
- Watts, D. (2007). A twenty-first century science. *Nature*, 445, 489. <https://doi.org/10.1038/445489a>
- Wessel, F., Schluter, R., Macherey, K., & Ney, H. (2001). Confidence measures for large vocabulary continuous speech recognition. *IEEE Transactions on speech and audio processing*, 9(3), 288-298.
- World Bank (2005). *Review of World Bank Conditionality*. Washington DC: The World Bank Group.
- World Bank (2011). *Education Sector Strategy 2020*. Washington DC: The World Bank Group.

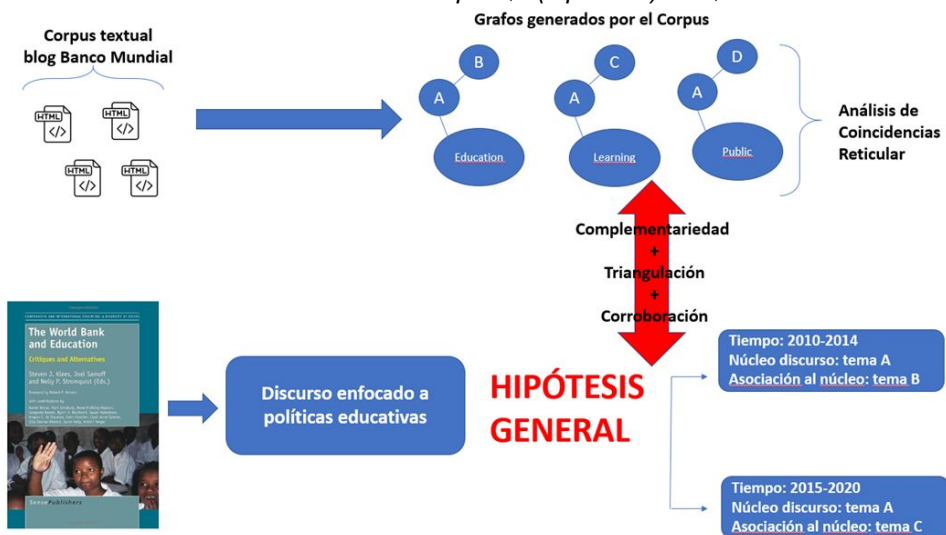
ANEXO

Ilustración 1. Estructura general del diseño de análisis



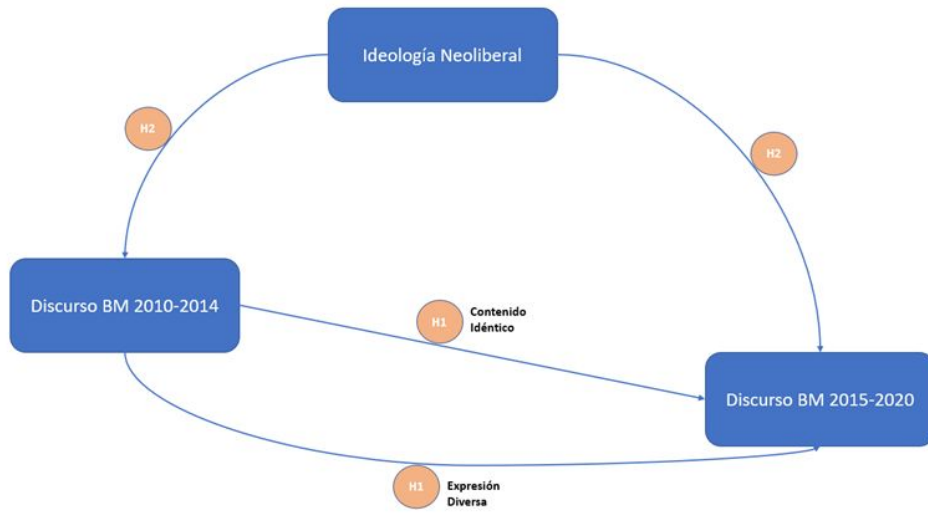
Fuente: Elaboración propia

Ilustración 2. Detalle de diseño de análisis de etapa QL (hipótesis) -> QN



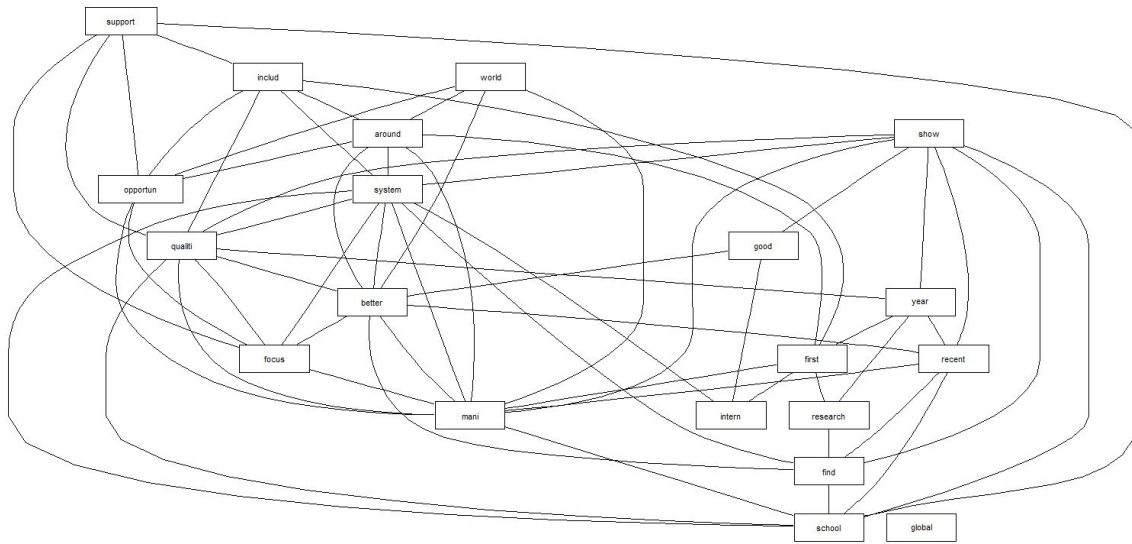
Fuente: Elaboración propia

Ilustración 3. Modelo de Análisis



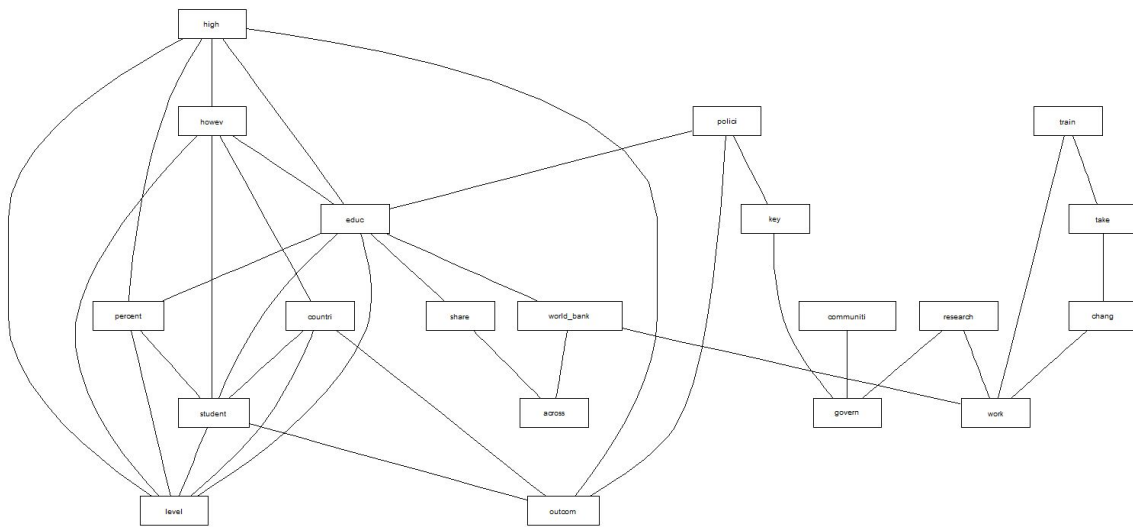
Fuente: Elaboración propia

Gráfico 1. Grafo de palabras para el período 2010-2014



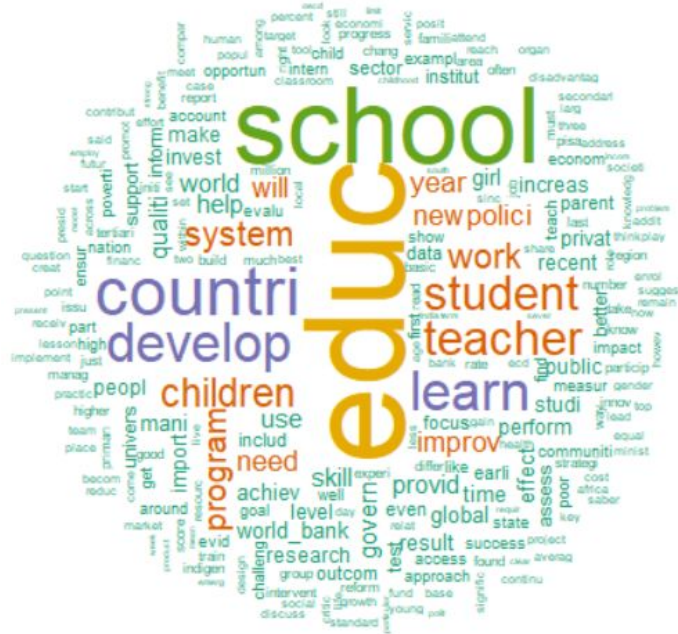
Fuente: Elaboración propia

Gráfico 2. Grafo de palabras para el período 2015-2020



Fuente: Elaboración propia

Gráfico 3. Cloud of words período 2010-2014



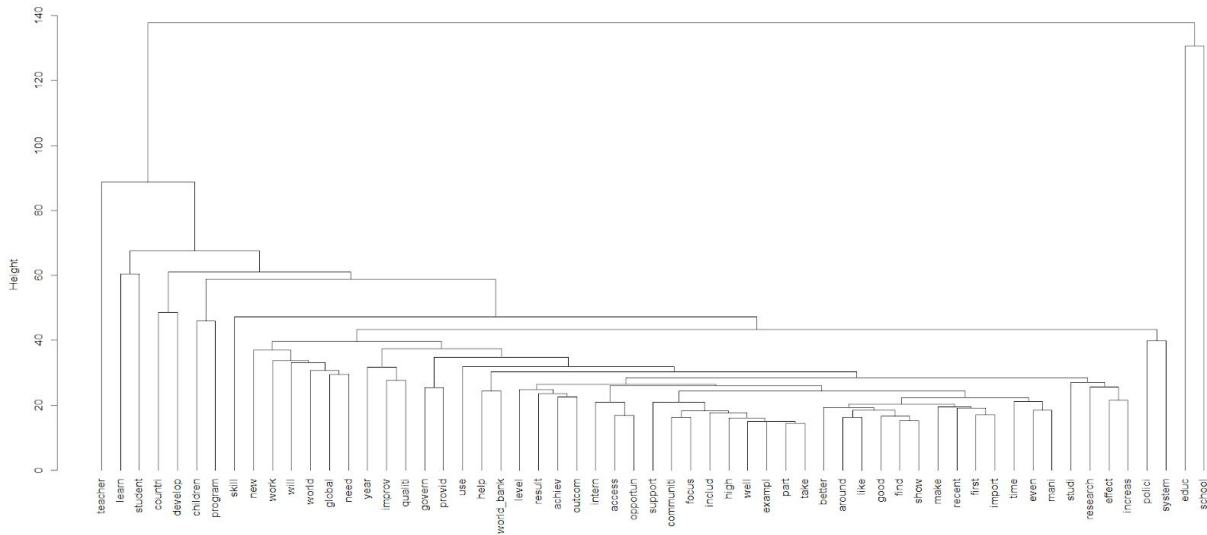
Fuente: Elaboración propia

Gráfico 4. Cloud of words período 2015-2020



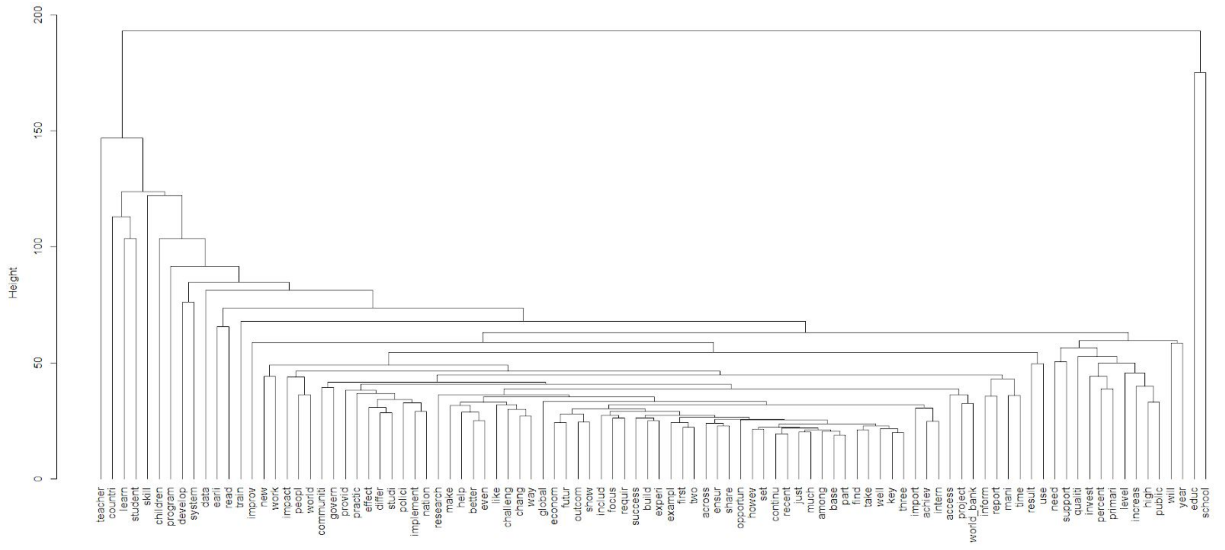
Fuente: Elaboración propia

Gráfico 5. Dendograma para período 2010-2014



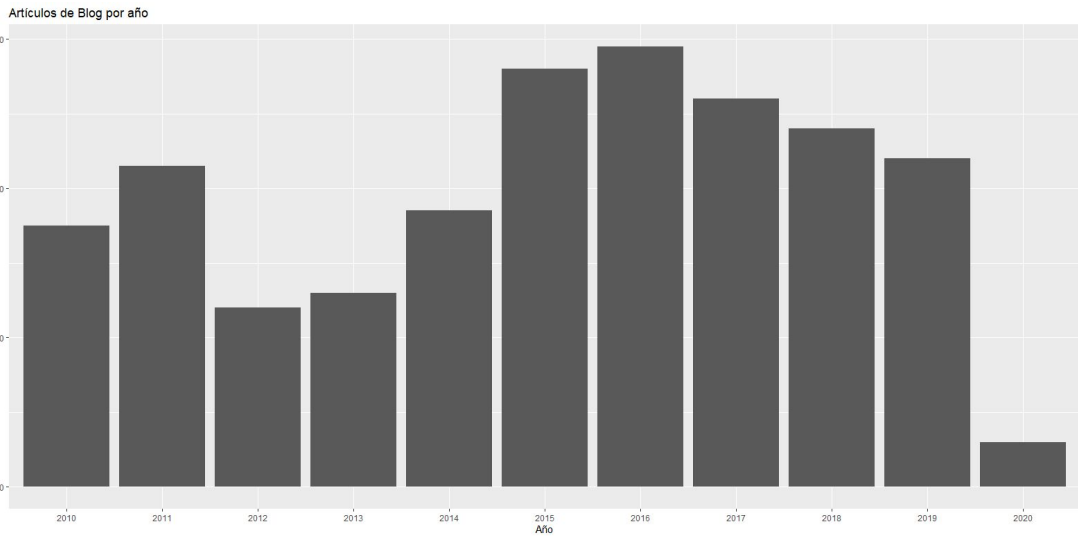
Fuente: Elaboración propia

Gráfico 6. Dendograma para período 2015-2020



Fuente: Elaboración propia

Gráfico 7. Cantidad de Artículos del Blog “Education for Global Development” agrupado por Año de publicación



Fuente: Elaboración propia

Tabla I. Palabras de mayor coeficiente de inclusión mutua respecto a la palabra clave “Education”, ambos períodos

Educ			
2010-2014		2015-2020	
Palabra	CIM	Palabra	CIM
<i>access</i>	0.39	<i>higher</i>	0.37
<i>financ</i>	0.39	<i>equiti</i>	0.34
<i>inequ</i>	0.34	<i>inclus</i>	0.32
<i>opportun</i>	0.34	<i>system</i>	0.28
<i>qualit</i>	0.33	<i>tertiari</i>	0.28

Fuente: Elaboración propia

Tabla II. Palabras de mayor coeficiente de inclusión mutua respecto a la palabra clave “Learning”, ambos períodos

Learn			
2010-2014		2015-2020	
Palabra	CIM	Palabra	CIM
<i>measur</i>	0.52	<i>crisi</i>	0.4
<i>focus</i>	0.44	<i>scatter</i>	0.34
<i>smarter</i>	0.44	<i>meritocrat</i>	0.33
<i>slum</i>	0.41	<i>text</i>	0.3
<i>highlevel</i>	0.39	<i>decent</i>	0.29

Fuente: Elaboración propia

Tabla III. Palabras de mayor coeficiente de inclusión mutua respecto a la palabra clave “Public”, ambos períodos

Public			
2010-2014		2015-2020	
Palabra	CIM	Palabra	CIM
<i>concess</i>	0.63	<i>relief</i>	0.46
<i>privat</i>	0.62	<i>opposit</i>	0.45
<i>scarc</i>	0.52	<i>examin</i>	0.42
<i>provid</i>	0.49	<i>candid</i>	0.41
<i>school</i>	0.47	<i>shadow</i>	0.4

Fuente: Elaboración propia

Tabla IV. Palabras de mayor coeficiente de inclusión mutua respecto a la palabra clave “Private”, ambos períodos

Private			
2010-2014		2015-2020	
Palabra	CIM	Palabra	CIM
<i>provid</i>	0.74	<i>return</i>	0.52
<i>sector</i>	0.73	<i>incom</i>	0.5
<i>public</i>	0.62	<i>contract</i>	0.49
<i>clearer</i>	0.61	<i>justifi</i>	0.46
<i>regulatori</i>	0.52	<i>higher</i>	0.45

Fuente: Elaboración propia

Tabla V. Palabras de mayor coeficiente de inclusión mutua respecto a la palabra clave “Teacher”, ambos períodos

Teacher			
2010-2014		2015-2020	
Palabra	CIM	Palabra	CIM
<i>student</i>	0.43	<i>effect</i>	0.36
<i>effect</i>	0.39	<i>train</i>	0.35
<i>time</i>	0.3	<i>practic</i>	0.3
<i>use</i>	0.29	<i>learn</i>	0.29
<i>system</i>	0.27	<i>system</i>	0.24

Fuente: Elaboración propia

Tabla VI. Palabras de mayor coeficiente de inclusión mutua respecto a la palabra clave “Student”, ambos períodos

Student			
2010-2014		2015-2020	
Palabra	CIM	Palabra	CIM
<i>score</i>	0.44	<i>firstyear</i>	0.38
<i>classroom</i>	0.43	<i>academ</i>	0.34
<i>perform</i>	0.43	<i>tuition</i>	0.33
<i>teacher</i>	0.43	<i>fundrais</i>	0.3
<i>municip</i>	0.41	<i>gpa</i>	0.3

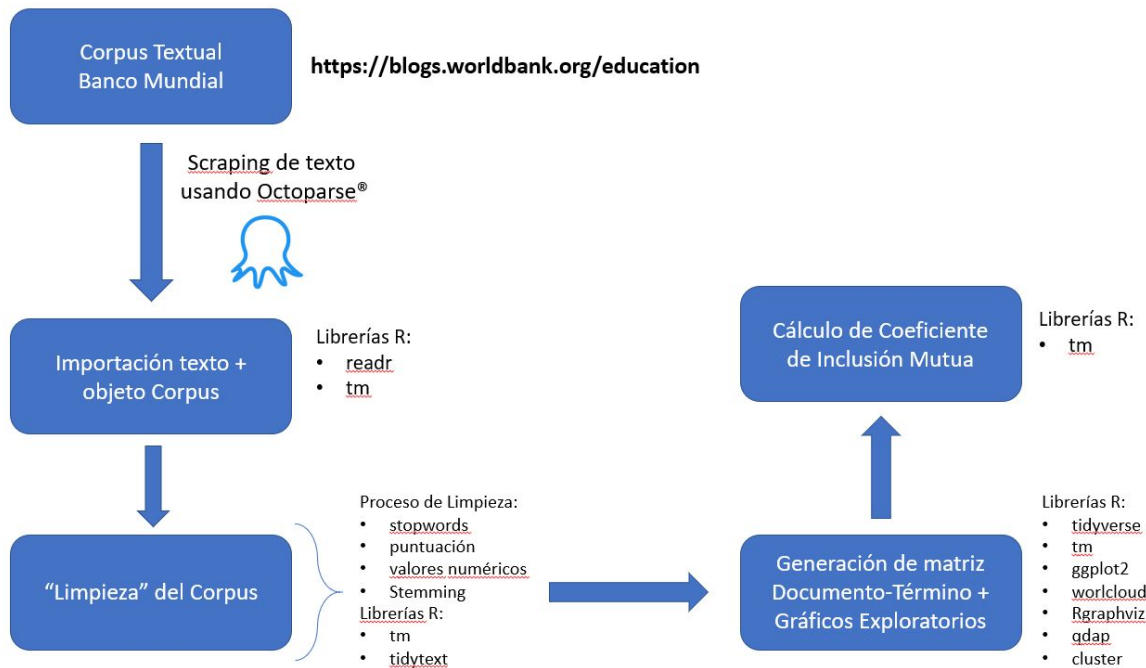
Fuente: Elaboración propia

Tabla VII. Indicadores y Densidad de los grafos analizados

		2010-2014	2015-2020
<i>g</i>	# nodos	20	20
<i>L</i>	# links	55	34
$2L / (g * (g-1))$	Densidad	0.2895	0.1789

Fuente: Elaboración propia

Circuito de Procesamiento 1. Circuito de recogida y procesamiento de datos, y cálculo de indicadores y gráficos utilizados en el análisis. El circuito sigue los pasos del script disponible en DOI: <https://doi.org/10.5281/zenodo.3836126>



Fuente: *Elaboración propia*