

Identificación de áreas temáticas científicas mediante el análisis de co-palabras. El caso de la Universidad de La Frontera, Chile¹

Patricio Padilla
José Coloma
Jorge Petit Breuilh*
Leonardo Mena
Daniel Barrios
Daniel Weinacker
Fabiola Vásquez
Blanca Villalobos**
Cristian Alister
Ronald Cancino***

Artículo recibido:
16 de octubre de 2014.

Artículo aceptado:
11 de marzo de 2016.

RESUMEN

El artículo presenta una experiencia de análisis y gestión de información científica y tecnológica para la focalización de la Investigación, Desarrollo e Innovación (I+D+i) en la Universidad de La Frontera, Chile. El objetivo fue identificar y visualizar áreas temáticas de la producción científica de la universidad visibles en el sistema de publicación científica y en el sistema de financiamiento público. La muestra consideró publicaciones y proyectos

* Instituto de Desarrollo Local y Regional. Universidad de la Frontera, Chile.

patricio.padilla@ufrontera.cl jr.colomaz@gmail.com jorge.petit@ufrontera.cl

** Instituto de Desarrollo Local y Regional. Dirección de Innovación y Transferencia Tecnológica, Universidad de La Frontera, Chile.

leomenac@gmail.com daniel.barrios@ufrontera.cl daniel.weinacker@gmail.com
fabiola.vasquez@ufrontera.cl blanca.villalobos@ufrontera.cl

*** Instituto de Desarrollo Local y Regional. Departamento de Ciencias Sociales, Universidad de la Frontera, Chile. calister@uc.cl roald.cancino@ufrontera.cl

disponibles durante el periodo 2007-2011. La identificación de las áreas temáticas se aborda desde la utilización complementaria del método de las palabras asociadas y la validación de expertos, mediante el análisis temático de contextos elementales y la utilización de herramientas de visualización de redes sociales. Se extraen aprendizajes de gestión de información y gestión de capacidades científicas y tecnológicas.

Palabras clave: Áreas Temáticas; Bibliometría; Análisis de Co-palabras; Chile.

ABSTRACT

Identification of scientific topic areas through co-words analysis. A case study from the Universidad de La Frontera, Chile

Patricio Padilla, José Coloma, Jorge Petit-Breuilb, Leonardo Mena, Daniel Barrios, Daniel Weinacker, Fabiola Vásquez, Blanca Villalobos, Cristian Alister and Ronald Cancino

This paper presents an experience of the analysis and management of scientific and technological information for directing research, development and innovation efforts in the Universidad de La Frontera. The study's objective is to identify and envision research fields in the University's scientific production appearing in published scientific research and recorded in the public financing systems. The sample includes publications and projects published between 2007 and 2011. The task of research field identification is approached through the complementary use of "Co-words Analysis"; "Expert Validation Methodology" and "Thematic Analysis of Elementary Contexts"; in conjunction with social network visualization tools. Information regarding the management of information and scientific and technological capabilities are among the study's findings.

Keywords: Thematic Areas; Bibliometrics; Co-Analysis-Co-Words Analysis; Chile.

INTRODUCCIÓN

En el año 2006 la Universidad de La Frontera (UFRO) crea su Vicerrectoría de Investigación y Postgrado, lo que permitió aumentar significativamente los proyectos de investigación básica y aplicada adjudicados, así como de publicaciones y patentes solicitadas. Además, con la creación de la Dirección de Innovación y Transferencia Tecnológica (DITT) en ese mismo año, se incorporó la ejecución de proyectos de alto impacto económico y social que vinculan a la universidad con entidades públicas y privadas mediante la investigación aplicada al quehacer universitario. A partir de los logros alcanzados en la formulación, gestión y seguimiento de proyectos de investigación aplicada, se crea además la Unidad de Transferencia Tecnológica (UTT) en 2009, una unidad adicional a la Unidad de Proyectos de I+D+i y a la Unidad de Gestión y Apoyo a investigadores que existían previamente en la DITT. Desde el punto de vista de los resultados obtenidos a partir de estos cambios, en la actualidad la UFRO tiene en marcha 35 proyectos de I+D que involucran aproximadamente 55 entidades asociadas, entre empresas y otras organizaciones, que le han permitido en pocos años tener cerca de 28 solicitudes de patentes en curso, 19 patentes ya concedidas, 12 marcas registradas para distintos productos de estas Tecnologías UFRO, dos marcas en proceso de registro, 18 protecciones de software y ocho manuales científicos, a través de derecho de autor. La UFRO posee además cinco licencias y acuerdos de entendimiento con terceros en distintas áreas como software, medicina e ingeniería.

La DITT se encuentra desarrollando una estructura de soporte para el funcionamiento de la UTT que requirió identificar las líneas de investigación importantes de la universidad y así concentrar esfuerzos en las oportunidades de desarrollo de investigaciones con potencial de convertirse en prototipos de productos o invenciones escalables al mercado. En este contexto, se presentan dos experiencias respecto a la utilización complementaria del análisis de co-palabras y la validación de expertos para la identificación y visualización de áreas temáticas en ciencia y tecnología. La razón de incorporar las ciencias sociales se debió a la menor vinculación de esta área científica con procesos de transferencia tecnológica para el caso de la Universidad de La Frontera.

El análisis se abocó a la universidad en su conjunto y se especificó una dimensión a las ciencias sociales. Para los fines de la investigación, las “áreas temáticas” se entienden como la especialización y acoplamiento de los investigadores en tópicos disciplinares específicos.

En la primera sección se presenta el enfoque del método de las palabras asociadas, atendiendo a su relación con otros métodos de la cienciometría.

Posteriormente se detalla la metodología y se muestran los resultados. En el apartado final se realiza una discusión y se extraen las principales conclusiones.

EL ANÁLISIS DE CO-PALABRAS EN LA VISUALIZACIÓN DE ÁREAS TEMÁTICAS EN CIENCIA Y TECNOLOGÍA

El documento científico y técnico condensa intenciones y expectativas de comunicación (Maltrás y Quintanilla, 1996) tanto de los autores (títulos, resúmenes y palabras clave) como de las instituciones de comunicación hacia la comunidad científica (selección de documentos por pertinencia temática y expectativa de calidad). La cienciometría se encarga de estudiar este tipo de documentos, ya sea los recursos y resultados, como las formas de organización en la producción de conocimientos y técnicas (Callon, Courtial y Penan, 1995) expresadas en artículos científicos y patentes (Callon, Courtial y Penan, 1995; Arencibia y De Moya, 2008). La cienciometría ocupa indicadores bibliométricos para su cometido, es decir, aplica “medidas basadas habitualmente en recuentos de publicaciones, con las que se pretende cuantificar los resultados científicos atribuibles bien a unos agentes determinados o bien a agregados significativos de estos agentes” (Maltrás y Quintanilla, 1996: 109).

A lo menos tres tipos de indicadores bibliométricos son ampliamente utilizados (Maltrás y Quintanilla, 1996): i) los indicadores de actividad, que buscan cuantificar los resultados de la actividad científica; ii) los indicadores relacionales de primera generación, centrados en las redes que sostienen el ciclo de producción de esos resultados, y iii) los indicadores relacionales de segunda generación, que ingresan al contenido de los documentos científicos para dar cuenta de las expresiones y relaciones lógicas (Escalante, 2009).

Una cuestión central en este estudio es el modo en que se identifican las áreas temáticas. Una primera forma será denominada “clasificación manual” (Arnold y Haefner, 1988), donde el observador infiere relaciones temáticas a medida que analiza los documentos, lo que es un ejercicio altamente dependiente de los procesos cognitivos del observador, como por ejemplo la acumulación de saber experto referente al campo estudiado y, además, dependiente de la cantidad de documentos ya que a medida que aumenten en número, el ejercicio se vuelve complejo, sobre todo si la pretensión es analizar la producción total de una institución o disciplina. Una segunda forma es la “clasificación mediada por computadora”, que incluye estudios enmarcados en los indicadores relacionales de segunda generación, cuyas técnicas se centran en la utilización de herramientas informacionales que permiten analizar gran cantidad de información a través de algoritmos aplicados a la

información textual. Este último es el que desarrollaremos de manera más extensa.

El método de las palabras asociadas o co-palabras y su visualización en mapas semánticos (Leydesdorff y Welbers, 2011) es parte de los indicadores relacionales de segunda generación (Callon, Courtial y Penan, 1995; Jurado y Bailón, 2002). Los pioneros en la utilización del análisis de co-palabras fueron Rip y Courtial (1984), quienes estudiaron los artículos generados en 10 años de una revista especializada en biotecnología y relacionaron las publicaciones y la variedad en mapas de biotecnología analizando sus conexiones de actualidad a través de los años. Posteriormente, en 1991 se genera el emblemático estudio de Michael Callon sobre interacciones disciplinares entre distintos tipos. La investigación se apoyó en el análisis de contenido en información sobre ciencia en polímeros, evidenciando que algunas áreas temáticas fueron generadas inicialmente en programadores académicos y otros por secciones de investigación tecnológicas. Uno de los grandes beneficios de este trabajo fue considerar el análisis de co-palabras como un método que podría aumentar la interacción de la academia con la industria (Callon, Courtial y Laville, 1991).

En la década de los noventa destaca la flexibilidad del análisis de co-palabras y también la evidencia de algunas complicaciones. Mostró buenos resultados como método al ser sometido a prueba de escala y amplitud temática, permitiendo analizar el campo de la acidificación, pero considerando seis disciplinas que abarcaron más de 18 000 publicaciones entre 1980 y 1987 (Law y Whittaker, 1992). También ha permitido conocer espacios interdisciplinarios, específicamente en los trabajos que convergen en la teoría del actor-red (Courtial, 1994). Sin embargo, en un emblemático estudio que intentó descubrir patrones de valorización industrial de proteínas vegetales apoyado en distintas bases de datos, evidenció que se requería nutrir de otras técnicas de la ciencimetría para mejorar su nivel de confianza (De Looze y Lemarié, 1997), además de la posibilidad de integrar a un panel de expertos en la interpretación de información (Bhattacharya y Basu, 1998), realizado por Lee y Jeong (2008) en el área de la robótica coreana que analizó metadatos de proyectos nacionales.

Recientemente, Van den Besselaar y Heimeriks (2006) señalan que los temas de investigación pueden ser representados mediante dos factores, extracción de palabras clave en los títulos de los artículos y el rango de publicaciones mediante citas, lo que permitiría mayor cobertura en el estudio de la literatura científica. En 2011 un estudio abordó la estructura de atención en el campo de los cuidados paliativos mediante el método de palabras asociadas complementado con minería de texto y redes neuronales; sus resultados

mostraron que las ciencias de la salud estaban fuertemente vinculadas a la investigación y servicios en el área de los cuidados paliativos (Yang, Bhikshu y Tsaih, 2011). En América Latina existen experiencias relevantes para el presente estudio, como la identificación de temáticas de investigación y características de la diseminación de información en una revista científica (Ronda *et al.*, 2013) y la combinación de técnicas de la lexicometría con el análisis de redes sociales para visualizar mapas temáticos, específicamente para bibliotecas y sus distintas colecciones por materias (Miguel, Caprile y Jorquera, 2008).

Estos estudios han permitido no sólo conocer la emergencia de áreas de investigación, sino también conformar comunidades científicas, vincular los niveles de dispersión de determinadas áreas temáticas respecto al comportamiento de estas propias comunidades científicas, así como también proveer de información a políticas de fomento a la ciencia, tecnología e innovación.

Respecto a las herramientas que facilitan el método de palabras asociadas, existen diversos software especializados para el tratamiento de datos textuales que permiten procesar grandes volúmenes de información documental y bibliográfica, reduciendo mediante algoritmos su complejidad. Uno de ellos es T-LAB, software compuesto por un conjunto de herramientas lingüísticas y estadísticas para el análisis de contenido, el análisis del discurso y la minería de textos.

Los *outputs* que el software dispone (tablas y gráficos) son un tipo particular de textos que pueden ser interpretados como mapas de isotopías (*iso* = igual; *topos* = lugar), es decir, como mapas de temas “genéricos” o “específicos” (Rastier, Cavazza y Abeillé, 2002) compuestos por la co-ocurrencia de componentes semánticos. Técnicamente son objetos multi-semióticos caracterizados por el hecho de que las relaciones entre los signos y los símbolos están ordenados por medidas que redireccionan a códigos específicos (Lancia, 2012), por lo que posibilitan realizar inferencias sobre las relaciones que organizan las formas del contenido, debido a que las relaciones de las formas de expresión no son casuales, sino que están constituidas por las medidas y algoritmos que determinan la organización de los espacios semánticos multidimensionales (Lancia, 2007). Su uso se extiende desde el análisis de la literatura científica y técnica para representar los conocimientos sobre áreas temáticas determinadas (Balzano *et al.*, 2009; Capone y Petrillo, 2011; Kirkels, 2012), incluyendo estudios de co-palabras sobre educación (Assefa y Rorissa, 2013), hasta las formas que adquieren el discurso, las representaciones sociales y prácticas reflexivas en diversos grupos humanos (Benavente y Cortini, 2010; Bruno, Galuppo y Gilardi, 2011; De Rosa y Holman, 2011) y en la web social (Veltri, 2013).

En este contexto, la identificación de áreas temáticas permite mejorar las capacidades de gestión de la investigación, desarrollo e innovación puesto que logra traducir a ojos de los gestores las capacidades de conocimiento tácito (know-how) acumulado en áreas específicas de desarrollo científico tecnológico que han sido priorizadas por la Universidad de La Frontera, según el Plan Estratégico de Desarrollo 2013-2023 y la detección de áreas emergentes no priorizadas que podrían tener alto impacto. Todo ello permite una focalización de recursos de transferencia tecnológica, obteniendo insumos para el desarrollo de investigaciones aplicadas conformes al plan señalado anteriormente, que ha definido la existencia de un monitoreo constante del impacto de la investigación realizada por la universidad. Se trata de proyectar escenarios futuros que puedan afectar la ubicación relativa de la institución en el ámbito nacional e internacional. Lo anterior refiere a los indicadores de productividad, a la conformación de bases para el establecimiento de redes con grupos de investigación con productividad equivalente y al fortalecimiento de las capacidades existentes que aseguren la transferencia tecnológica efectiva del conocimiento producido en la universidad, contribuyendo a generar una cultura de innovación y emprendimiento.

El método de las palabras asociadas permite hacer caracterizaciones cognitivas de distintas áreas del conocimiento y ha sido una herramienta de gran ayuda para la gestión científica en centros de investigación y especialmente en universidades. Sin embargo, posee limitaciones a la hora de explicar la forma en que ciertas áreas se cristalizan en conglomerados temáticos, es decir, permitiría una visualización y la forma en que se estructura el conocimiento, pero con la necesidad de ser complementado con técnicas estadísticas y organizacionales que fundamenten dicha estructuración cognitiva. Las dificultades del análisis de co-palabras han sido tempranamente advertidas por Law y Whittaker (1992) debido a la generalidad de los indizadores que incluso en ocasiones pueden tener baja pertinencia. Para ello, las recomendaciones consisten en someter el trabajo de corpus a procesos de normalización y realizar análisis en apoyo de distintos expertos de las disciplinas en cuestión (Wang *et al.*, 2012).

METODOLOGÍA

Se propone una aproximación metodológica enfocada en la utilización del análisis temático de contextos elementales (del software T-LAB) y el uso complementario de herramientas propias del análisis de redes sociales para la visualización de las co-ocurrencias de palabras, que sirvan de insumo

para la identificación de áreas temáticas científicas mediante la validación de expertos. La población está compuesta por los artículos científicos de la colección principal de Web of Science (WOS), la plataforma SciELO y los Proyectos del Fondo de Investigación Científica y Tecnológica en Chile, FONDECYT, provenientes de la Comisión Nacional de Investigación Científica y Tecnológica en Chile.

A partir de ello se construyeron dos casos pertenecientes a la producción científica de la Universidad de La Frontera:

Caso 1: comprende todos los artículos WOS y SciELO de la Universidad de La Frontera publicados durante el periodo 2007-2011.

En este caso la mayor parte de los artículos estaban en inglés, por lo que los documentos en español fueron traducidos para el análisis.

Tabla 1. Publicaciones por año de la Universidad de La Frontera

Año	WOS	SciELO
2007	95	72
2008	139	97
2009	162	108
2010	215	128
2011	233	165
Total	844	570

Caso 2: comprende los artículos WOS, SciELO y los Proyectos FONDECYT adjudicados durante el periodo 2007-2011, enmarcados en el área de las Ciencias Sociales de la Universidad de La Frontera. Como se señaló en la introducción, el énfasis en ciencias sociales se debe a que esta área científica presenta menores niveles de transferencia tecnológica, por lo que la información se complementó con proyectos para obtener mayor profundidad en el análisis.

A diferencia del caso anterior, la mayor parte de los artículos estaban en español, por lo que se optó por traducir los escasos artículos en inglés.

Tabla 2. Publicaciones y proyectos por año de la Universidad de La Frontera

Año	Proyectos	Publicaciones
2007	2	13
2008	9	24
2009	7	17
2010	6	18

2011	3	24
Total	27	96

Ambas tablas contienen el total de documentos sin duplicados. Cuando un documento presentaba indexación SciELO y WOS simultáneamente, se optó por clasificarlo en este último asumiendo que esta dimensión no es fundamental para el estudio. A su vez, de ambos casos se extrajeron los títulos, resúmenes y palabras clave de los documentos con el objetivo de generar una base de datos.

Posteriormente, se prosiguió con la depuración de los datos consistente en i) eliminar espacios, ii) eliminar información de documentos científicos repetida y iii) transformar todas las palabras en minúscula.

Considerando las sugerencias de la literatura, el análisis de la información incorporó a expertos de la universidad, específicamente de los directores de investigación y postgrado pertenecientes a todas las facultades, que además asesoran a la Dirección de Investigación de la Universidad de La Frontera. Su rol fue validar el análisis mediante distintas etapas encaminadas a afinar el método y fuentes utilizadas, corregir aporías de la información levantada, precisiones temáticas y finalmente la interpretación conjunta de las áreas temáticas.

Análisis temático de contextos elementales

Luego de un proceso de preparación, importación y homogenización lexical de los corpus (títulos y resúmenes de las publicaciones y proyectos seleccionados), esta herramienta de análisis permite obtener una representación de los contenidos del corpus mediante pocos y significativos cluster temáticos que corresponden a un conjunto de títulos y resúmenes determinados por los mismos patrones de palabras clave, pudiendo ser descritos a partir de sus lemas (palabras que han sido lematizadas) y características.

Las palabras clave se obtuvieron a través de la modalidad de configuración personalizada disponible en el software. Los criterios utilizados fueron los siguientes:

- a) Relevancia: a través de la verificación cuantitativa (cantidad de ocurrencias) de las palabras.
- b) Significancia: a través de la verificación cualitativa de la no banalidad del significado de las palabras.

El proceso de análisis como tal se compone de las siguientes etapas:

- a) construcción de una tabla de unidades de contexto (título-resumen) x unidades lexicales (palabras clave) con valores de tipo presencia-ausencia;
- b) cálculo de pesos TF-IDF y normalización de los vectores (norma euclídea);
- c) clusterización de las unidades de contexto (medida de semejanza: coeficiente del coseno; método de clusterización: bisecting K-means);
- d) salvaguardia de las particiones obtenidas y para cada una de ellas;
- e) construcción de una tabla de contingencia unidades lexicales x clusters;
- f) test del chi cuadrado aplicado a todos los cruces unidades lexicales x clusters, y
- g) análisis de las correspondencias de la tabla de contingencia.

Visualización de co-ocurrencias de palabras clave o lemas

Las co-ocurrencias son el número de veces que dos o más palabras clave (lemas) están presentes contemporáneamente en los mismos títulos-resúmenes de un artículo científico o proyecto. El análisis permite observar, para cada lema, el conjunto de palabras que aparecen reiteradamente junto a ella, independiente a los cluster temáticos que pertenezca. Por tanto, la visualización de sus co-ocurrencias de palabras favorece una mayor descripción e interpretación de los contenidos tanto en su totalidad (conjunto de temas) como en específico (por cluster temático), que potencian el trabajo en talleres con expertos para facilitar la gestión científica en la universidad.

Debido a que T-LAB sólo visualiza las co-ocurrencias por cada lema, fue necesario exportar los datos con el objeto de visualizar simultáneamente las redes de ocurrencias de lemas a través del software NetDraw (Borgatti, 2002).

Finalmente, estos mapas de co-palabras sirvieron de insumos para la identificación de áreas temáticas en dos sentidos, correspondientes a los casos elegidos:

- a) Caso 1: análisis enfocado a la Universidad de La Frontera en su conjunto, se identificaron macro áreas temáticas desarrolladas por la universidad, a partir de la totalidad de las redes de ocurrencia de lemas, que resultaron ser áreas disciplinares según la clasificación OCDE. La clasificación OCDE se consideró como un criterio de

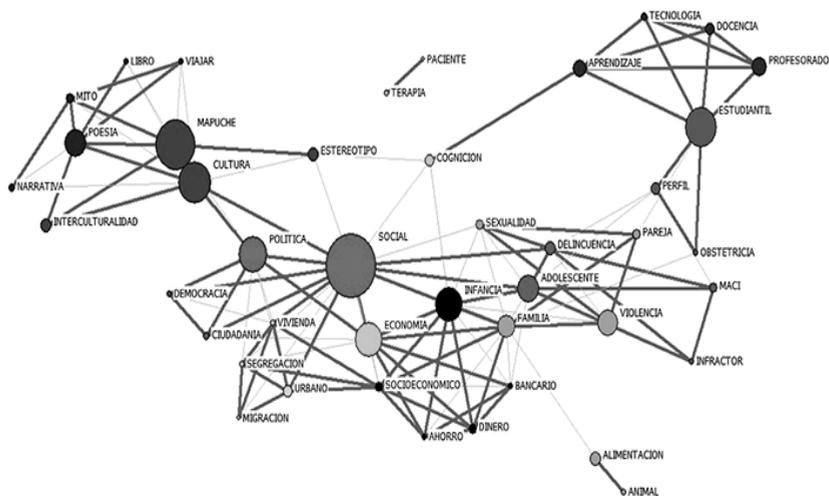


Gráfico 2. Áreas temáticas Ciencias Sociales

En Ciencias Médicas las áreas más representativas son medicina básica y clínica. En medicina básica los estudios con mayores capacidades instaladas corresponden al área de reproducción, específicamente en lo que concierne a espermatozoides: vitrificación, semen, incubación, daño en la membrana y mitocondria, entre otros. En medicina clínica los estudios se han concentrado principalmente en enfermedades y virus, como el virus del papiloma humano y el cáncer (colon, útero, vesícula y gástrico). Otra área clínica que se ha posicionado es la investigación odontológica, abarcando áreas referentes a tejido, maxilofacial, mandíbula, tensión, hueso, dientes, caries y cirugía, entre otros. Además, se han analizado variables externas, como fumar y el estrés, que podrían potenciar el daño producido tanto en la mandíbula como en los dientes.

En las Ciencias Agrícolas se encuentra la agricultura, silvicultura, pesca y ciencias afines, áreas que abarcan temas tan diversos como el estudio en frutas con propiedades antioxidantes (murta y el arándano), el cultivo de papa, flores y del suelo. En los estudios de frutas resaltan los flavonoides, polifenoles y antimicrobianos. En cuanto a la investigación de suelos –una de las líneas más importantes– destaca el estudio de suelos volcánicos, fertilizantes, biorremediación, desechos y el uso de bacterias, en especial las cianobacterias y rizobacterias.

En Ciencias Naturales se encuentran específicamente las matemáticas e informática con el uso de algoritmos para temas de optimización y el desarrollo

de programas atmosféricos. También ha servido como un incipiente apoyo a las ciencias sociales, considerando la vinculación entre lemas.

Respecto a las ciencias sociales, existe un grupo de ciencias de la educación y otras ciencias sociales. En las primeras se han explorado nuevas metodologías de enseñanza, utilizando distintas tecnologías de información y conocimiento aplicadas a la educación, cuya aplicación se ha realizado en escuelas rurales y colegios estatales de la región. En otras ciencias sociales, el concepto mapuche ha sido abordado desde una perspectiva social y agroindustrial. La perspectiva social incluye aspectos ligados a la cultura, discriminación y políticas públicas, mientras que la agroindustrial refiere al comercio y la utilización de ciertos alimentos por una cultura determinada: productos típicos y atractivos turísticos. Lemas de esta perspectiva son ingredientes, carne, fibra, mercado y arroz.

Específicamente, las áreas temáticas vinculadas a las Ciencias Sociales se dividen en 11 conglomerados que mantienen ciertos niveles de integración.

Tabla 3. Áreas temáticas en Ciencias Sociales de la Universidad de La Frontera

Temas	Palabras clave
Conglomerado 1	Mapuche, cultura, estereotipo e interculturalidad
Conglomerado 2	Adolescente, delincuencia, MACI e infractor
Conglomerado 3	Aprendizaje, docencia, y tecnología
Conglomerado 4	Social, política, ciudadanía y democracia
Conglomerado 5	Economía, cognición, terapia y paciente
Conglomerado 6	Violencia, familia, sexualidad y pareja
Conglomerado 7	Animal y alimentación
Conglomerado 8	Poesía, mito, narrativa, viajar y libro
Conglomerado 9	Urbano, segregación, vivienda y migración
Conglomerado 10	Estudiantil, perfil y obstetricia
Conglomerado 11	Infancia, dinero, socioeconómico, ahorro y bancario

Fuente: elaboración propia

El conglomerado primero se vincula al pueblo mapuche con los siguientes lemas, su nivel de presencia entre paréntesis: mapuche (1 477), cultura (1 190), estereotipo (336) e interculturalidad (312). El segundo conglomerado refiere a intervenciones de las ciencias sociales, específicamente de la psicología en infractores de ley, los lemas son adolescente (744), delincuencia (333), MACI (202) e infractor (101). El tercer conglomerado refiere a herramientas pedagógicas y educación: aprendizaje (412), docencia (284) y tecnología (195). El cuarto se relaciona con sociología y ciencias políticas en temáticas de participación ciudadana: social (1 069), política (1 008), ciudadanía (194) y democracia (119); el quinto con psicología económica: economía (978), cognición (256), terapia (111) y paciente (95); el sexto conglomerado se subdivide en un grupo referido a trabajo social en contextos familiares adversos: violencia (659), familia (580), sexualidad (241) y pareja (213), y otro grupo vinculado a productos de consumo: animal (117) y alimentación (341). El octavo conglomerado refiere a humanidades: poesía (720), mito (217), narrativa (167), viajar (111) y libro (110). El conglomerado noveno alude a temáticas vinculadas con estudios de migraciones y urbanismo: urbano (251), segregación (150), vivienda (110) y migración (76). El décimo a estudios específicos sobre perfiles de estudiantes universitarios: estudiantil (1136), perfil (269) y obstetricia (31). Finalmente, el onceavo conglomerado es una especialización del conglomerado cinco, vinculado a psicología e económica: infancia (957), dinero (229), socioeconómico (212), ahorro (141) y bancario (135).

DISCUSIÓN Y CONCLUSIONES

El análisis y visualización de capacidades de investigación, expresadas en publicaciones científicas en contextos universitarios, es una herramienta fundamental tanto para comprender el modo como una universidad se transforma hacia una organización compleja de producción, transmisión y diseminación de conocimiento, como para focalizar las políticas y modelos de gestión de la I+D+i. En este contexto, el artículo puede ser comprendido como un esfuerzo para estandarizar herramientas de análisis y visualización en la senda descrita.

En este contexto, las áreas temáticas más desarrolladas y representativas en los artículos científicos y proyectos de investigación de la Universidad de La Frontera presentan desarrollos desiguales. Entre los campos científicos con mayor desarrollo (integridad semántica en el análisis) se encuentran las ciencias médicas, ciencias del suelo y ciencias biológicas. Los aspectos comunes entre dichas áreas científicas son el resultado de procesos de gestión

que fortalecen las capacidades de investigación. En específico, ello se traduce en i) vinculación a programas de doctorado, ii) disposición de investigadores con alta productividad científica, traducida en publicación de artículos de corriente principal y adjudicación de patentes y iii) apoyo de financiamiento a través de convenios de desempeño entre la Universidad de La Frontera y el Ministerio de Educación de Chile; y fuertes capacidades de atracción de fondos de investigación, desarrollo e innovación (I+D+i). Las oportunidades de transferencia de tecnologías resultantes de investigaciones se explican por la existencia de capacidades en transformación de biomasa y utilización de la misma. Estas líneas se relacionan con ciencias transversales de investigación como modelamiento matemático, genética y potenciales coadyudantes.

En las ciencias sociales se permitió observar niveles bajos de integridad semántica que concuerdan con experiencias respecto a la fragmentación interpretativa en campos científicos, donde las ciencias sociales, a excepción de la economía, experimentan menores niveles de integración que otras áreas científicas. Para el caso de la sociología existen explicaciones derivadas de estudios que señalan el mayor cuestionamiento de los sociólogos sobre las bases conceptuales de la disciplina (Varga, 2011). Para el resto de las ciencias sociales no existen estudios que propongan explicaciones de éste u otro tipo. Una hipótesis posible sobre ello en la Universidad de La Frontera es que las ciencias sociales presentan tradiciones disciplinares de diseminación de conocimiento en formas de publicación no indexadas, como congresos y libros. Asociado a ello, en estas disciplinas existen capacidades crecientes de atracción de fondos para investigación básica, y sólo recientemente articulación a instrumentos orientados a la investigación y desarrollo. Así, recién en las ciencias sociales se genera una “cultura” de publicaciones en revistas indexadas, proceso de reciente incentivo mediante instrumentos de fomento como el Convenio de Desempeño en Ciencias Sociales. Un aspecto relevante es la posibilidad de potenciar el desarrollo de investigación aplicada y posterior transferencia –vínculos con el entorno–, cuyo foco temático central radica en el conocimiento generado en torno a multiculturalidad y educación en contextos de vulnerabilidad y otros. Ello se hace comprensible pues se articula con la declaración del Plan Estratégico de Desarrollo vigente de la Universidad de La Frontera, relativo a consolidar la investigación realizada mediante la creación del Núcleo en Ciencias Sociales, para integrar y conectar distintas áreas del conocimiento en la universidad. En este sentido, un aspecto relevante que emerge del análisis es la capacidad de articulación de las capacidades de investigación entre ciencias sociales, educación y las áreas más desarrolladas, como ciencias médicas, ciencias del suelo y ciencias biológicas. Este

aspecto será relevante para generar un entramado interdisciplinario de investigación en la Universidad de La Frontera.

Los resultados de investigación muestran el potencial de este tipo de análisis para promover procesos de I+D+i. En la Universidad de La Frontera ello se expresa en dos procesos, cuyos desafíos fundamentales son articular esfuerzos.

Primero, con la identificación de las áreas temáticas develadas, se obtienen insumos para enfocar esfuerzos de profesionales de la DITT, a fin de continuar y profundizar la gestión del desarrollo de investigaciones básicas y aplicadas, incrementando el volumen y eficacia del proceso de transferencia de las tecnologías y otros productos de investigación de la I+D realizada por la universidad. Segundo, este esfuerzo está encaminado a dotar a la Universidad de La Frontera de una unidad de estudios de la ciencia y bibliometría, en sintonía con la propuesta de Torres y Jiménez (2013), para de este modo articular capacidades de investigación bibliométricas, de sociología de la ciencia, y de políticas de ciencia, tecnología e innovación disponibles en la universidad. Con ello, será posible poner a disposición de la gestión de la investigación el análisis y la visualización de capacidades de investigación, diseminación, uso e impactos sociales y científicos de la ciencia y tecnología desarrollada y potencialmente por desarrollar en la Universidad de La Frontera. El presente ejercicio de investigación refleja que las posibilidades para una unidad de bibliometría no se agotan en procesamiento y análisis de información de forma utilitarista a la universidad, sino también en la realización de investigaciones que permitan enriquecer el abanico de estudios generados en la Universidad de La Frontera, con investigaciones vinculadas a a cienciometría, bibliometría y análisis de co-palabras.

REFERENCIAS

- Arencibia, R. y F. de Moya. 2008. "La evaluación de la investigación científica: una aproximación teórica desde la cienciometría." *ACIMED* 17 (4).
- Arnold, M. y C. Haefner. 1988. *Análisis del contenido de las tesis de grado en la Licenciatura en Antropología de la Universidad de Chile en sus menciones de antropología social y cultural: 1977 a 1987*. Serie Documentos de Estudio Antropología Social. Santiago: Universidad de Chile.
- Assefa, S. y A. Rorissa. 2013. "A bibliometric mapping of the structure of STEM education using co-word analysis." *Journal of the American Society for Information Science and Technology* 64 (12).
- Balzano, R., D. Sepio, M. Guidi, V. Puro, E. Girardi y N. Orchi. 2009. "Representation of HIV/Aids and mental health co-morbidity in medical and social sciences literature." *European Psychiatry* 24 (1).

- Benavente, P. y M. Cortini. 2010. "Human resource strategic management in NPOs: An explorative study on managers' psychosocial training." *Journal of Workplace Learning* 22 (8): 508-521.
- Bhattacharya, S. y P. K. Basu. 1998. "Mapping a research area at the micro level using co - word analysis." *Scientometrics* 43 (3): 359-372.
- Borgatti, S. 2002. *NetDraw Software for Network Visualization*. Lexington, KY: Analytic Technologies.
- Bruno, A., L. Galuppo y S. Gilardi. 2011. "Evaluating the reflexive practices in a learning experience." *European Journal of Psychology of Education* 26 (4): 527-543.
- Callon, M., J. P. Courtial y F. Laville. 1991. "Co-word analysis as a tool for describing the network of interactions between basic and technological research: The case of polymer chemistry." *Scientometrics* 22 (1): 155-205.
- Callon, M., J.-P. Courtial y H. Penan. 1995. *Cienciometría. El estudio cuantitativo de la actividad científica: de la bibliometría a la vigilancia tecnológica*. Gijón: Trea.
- Capone, V. y G. Petrillo. 2011. "Health Promotion in International Documents: Strengths and Weaknesses from the Perspective of Community Empowerment." *Journal of Community & Applied Social Psychology* 23 (2): 98-114.
- Courtial, J. P. 1994. "A cword analysis of scientometrics." *Scientometrics* 31 (3): 251-260.
- De Looze, M. A. y J. Lemarié. 1997. "Corpus relevance through co-word analysis: An application to plant proteins." *Scientometrics* 39 (3): 267-280.
- De Rosa, A. y A. Holman. 2011. "Social representations of female-male beauty and aesthetic surgery: a cross-cultural analysis." *Temas em Psicologia* 19 (1).
- Escalante, E. 2009. "Una nota metodológica sobre los análisis cualitativos. El análisis de las relaciones entre los elementos: El análisis de las frecuencias y co-ocurrencias." *Theoria* 18 (1): 57-67.
- Jurado, E. y R. Bailón. 2002. *Evaluación a través del análisis de las palabras asociadas. Aplicación a la evaluación de la investigación científica y técnica*. España: Universidad de Granada, Biblioteca de Andalucía.
- Kirkels, A. 2012. "Discursive shifts in energy from biomass: A 30 year European overview." *Renewable and Sustainable Energy Reviews* 16 (6): 4105-4115.
- Lancia, F. 2007. *Word Co-occurrence and Similarity in Meaning. Some Methodological Issues*. <http://mytlab.com/wcsmeaning.pdf>
- Lancia, F. 2012. *T-LAB 8.0 - User's Manual*. <http://www.tlab.it/es/bibliography.php>
- Law, J. y J. Whittaker. 1992. "Mapping acidification research: A test of the co-word method." *Scientometrics* 23 (3): 417-461.
- Lee, B. y Y. I. Jeong. 2008. "Mapping Korea's national RyD domain of robot technology by using the cword analysis." *Scientometrics* 77 (1): 3-19.
- Leydesdorff, L. y K. Welbers. 2011. "The semantic mapping of words and co-words in contexts." *Journal of Informetrics* 5 (3): 469-475.
- Maltrás, B. y M. Quintanilla. 1996. "Los indicadores bibliométricos en el estudio de la ciencia. Fundamentos conceptuales y aplicación en política científica." Tesis doctoral, Universidad de Salamanca.
- Miguel, S., L. Caprile e I. Jorquera. 2008. "Análisis de co-términos y de redes sociales para la generación de mapas temáticos de la colección", presentado en *V Jornadas de Sociología de la UNLP*, La Plata, 10, 11 y 12 de diciembre. FAHCE. Mesa J 29

- Sociedad, información, bibliotecas y profesionales. Desafíos actuales. *Plan Estratégico de Desarrollo 2013-2023*. Temuco: Universidad de La Frontera. <http://humanidades.ufro.cl/index.php/noticias/1168-plan-estrategico-de-desarrollo-2013-2023>
- Rastier, F., M. Cavazza y A. Abeillé. 2002. *Semantics for Descriptions*. Stanford: CSLI.
- Rip, A. y J. P. Courtial. 1984. "Co-word maps of biotechnology: An example of cognitive scientometrics." *Scientometrics* 6 (6): 381-400.
- Ronda, G., C. Batista, J. C. Domínguez y R. Ronda. 2013. "Cambios en la estructura intelectual de la investigación sobre temas de dirección en Cuba: un análisis de co-palabras de Folletos Gerenciales 1997-2009." *Bibliotecas. Anales de Investigación* 8-9 (8-9):110-120.
- Torres, D. y E. Jiménez. 2013. "Hacia las unidades de bibliometría en las universidades: modelo y funciones." *Revista Española de Documentación Científica* 35 (3): 469-480.
- Van den Besselaar, P. y G. Heimeriks. 2006. "Mapping research topics using word-reference co-occurrences: A method and an exploratory case study." *Scientometrics* 68 (3): 377-393.
- Varga, A. 2011. "Measuring the semantic integrity of scientific fields: a method and a study of sociology, economics and biophysics." *Scientometrics* 88: 163-177.
- Veltri, G. A. 2013. "Microblogging and nanotweets: Nanotechnology on Twitter." *Public Understanding of Science* 22 (7): 832-849.
- Yang, Y.-H., H. Bhikshu y R.-H. Tsaih. 2011. "The topic analysis of hospice care research using co-word analysis and GHSOM." *Intelligent Computing and Information Science* 134: 459-465.
- Wang, Z.-Y., G. Li, C.-Y. Li y A. Li. 2012. "Research on the semantic-based co-word analysis." *Scientometrics* 90: 855-875.

Los autores agradecen al Ministerio de Educación de Chile y a la Universidad de La Frontera por apoyar la presente investigación a través del Plan de Mejoramiento Institucional: Modelo Interdisciplinario de Gestión Articuladora para el Desarrollo Humano Territorial Equilibrado en Comunas Piloto de la Región de La Araucanía. Convenio de Desempeño Regional, FRO 1301. Los autores también agradecen los valiosos comentarios de la evaluación anónima de la Revista Investigación Bibliotecológica.

Para citar este texto:

Padilla-Patricio, José Coloma, Jorge Petit-Breuilh, Leonardo Mena, Daniel Barrios, Daniel Weinacker, Fabiola Vásquez, Blanca Villalobos, Cristian Alister y Ronald Cancino. 2017. "Identificación de áreas temáticas científicas mediante el análisis de co-palabras. El caso de la Universidad de La Frontera, Chile". *Investigación Bibliotecológica: Archivonomía, Bibliotecología e Información* 72 (31): 63-80.
<http://dx.doi.org/10.22201/iibi.0187358xp.2017.72.57824>