



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
PROGRAMA DE MAESTRÍA Y DOCTORADO EN INGENIERÍA
SISTEMAS – PLANEACIÓN

ESTADO DEL ARTE DEL CAMPO DISCIPLINARIO DE LA EVALUACIÓN

TESIS
QUE PARA OPTAR POR EL GRADO DE:
MAESTRA EN INGENIERÍA

PRESENTA:
ING. MARÍA DE LOS ÁNGELES ZAMORA TORRES

TUTOR PRINCIPAL
GABRIEL DE LAS NIEVES SÁNCHEZ GUERRERO
PROGRAMA DE MAESTRÍA Y DOCTORADO EN INGENIERÍA

CIUDAD DE MÉXICO OCTUBRE 2020



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Jurado

- Presidente:** Dr. José Jesús Acosta Flores
- Secretario:** Dr. Javier Suárez Rocha
- Primer vocal:** Dr. Gabriel de las Nieves Sánchez Guerrero
- Segundo vocal:** Dr. José Antonio Rivera Colmenero
- Tercer vocal:** M en I. Francisca Irene Soler Anguiano

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

Tutor de tesis

Dr. Gabriel de las Nieves Sánchez Guerrero

Firma

AGRADECIMIENTOS

A la Universidad Nacional Autónoma de México por brindarme, desde pequeña hasta el día de hoy, la oportunidad de crecer personal y profesionalmente en su Ciudad Universitaria.

A mis padres por su apoyo incondicional en cada una de las decisiones que he tomado a lo largo de mi vida. Gracias por todo su amor.

A mis profesores de la maestría por su dedicación y esfuerzo. En especial al Dr. Gabriel de las Nieves Sánchez Guerrero por sus consejos y asesoría para la realización de este trabajo. Así mismo, a mis sinodales, la M. I. Francis Soler, el Dr. Javier Suárez Rocha, el Dr. Jesús Acosta y el Dr. José Antonio Colmenero por sus comentarios y observaciones a esta investigación.

A los ingenieros Andrés Mota y Rosalba Huerta, y Roberto Chávez de Grupo HMG por encaminarme a la vida profesional.

A mis amigas Pilar, Xochitl y Karina.

A mis amigos Omar, Abdula, Edher, Adán, Roberto (Homeworks), Angelito, Leonardo, José Miguel, Sergio Cruz y Pedro.

A mis amigos ciclistas Adrián, Noé, Leo y Osvaldo.

CONTENIDO

RESUMEN	1
ABSTRACT	2
INTRODUCCIÓN	3
ANTECEDENTES	5
OBJETIVOS	19
BÚSQUEDA, REFERENCIAS CONCEPTUALES Y FUENTES DE INFORMACIÓN	21
ANÁLISIS DE LA INFORMACIÓN	31
VISUALIZACIÓN Y DISCUSIÓN	35
Evaluación	38
Sistemas	62
Educación	74
CONCLUSIONES	79
BIBLIOGRAFÍA	82
ANEXO A	85
ANEXO B	105

ÍNDICE DE TABLAS

Tabla 1–Generaciones de la evaluación _____	17
Tabla 2–Enfoques de evaluación _____	18
Tabla 3 – Número de revistas por campo de conocimiento. _____	30

ÍNDICE DE MAPAS

Mapa 1. Mapa bibliométrico de densidad de cocitación de autores de las revistas de evaluación. _____	41
Mapa 2. Mapa de redes de cocitación de autores de las revistas de evaluación. _____	42
Mapa 3. Mapa de redes de coocurrencia de palabras claves de las revistas de evaluación con aristas. _____	44
Mapa 4. Mapa de redes de visualización superpuesta por año de publicación en el área de evaluación. _____	51
Mapa 5. Mapa de coautoría en el área de evaluación. _____	61
Mapa 6. Mapa de redes de coocurrencia de palabras claves de las revistas de sistemas y planeación con aristas. _____	63
Mapa 7. Mapa bibliométrico de densidad de coocurrencia de palabras claves de las revistas de sistemas y planeación. _____	64
Mapa 8. Mapa de red con visualización superpuesta de coocurrencia de las revistas de sistemas y planeación. _____	65
Mapa 9. Mapa de red con visualización superpuesta de coocurrencia de las revistas de sistemas y planeación. _____	67
Mapa 10. Mapa de red con visualización superpuesta de coocurrencia de las revistas de sistemas y planeación. _____	68
Mapa 11. Mapa de red con visualización superpuesta de coocurrencia de las revistas de sistemas y planeación. _____	71
Mapa 12. Mapa de red de coocurrencia de palabras claves de las revistas de educación. _____	75
Mapa 13. Mapa de red de coocurrencia de palabras con visualización superpuesta de las revistas de educación. _____	76

Mapa 14. Mapa de red de coocurrencia de palabras con visualización superpuesta de las revistas de educación.	77
Mapa 15. Clúster 1 (rojo). Developmental Evaluation.	86
Mapa 16. Clúster 2 (verde). Evaluation Practice.	87
Mapa 17. Clúster 3 (azul fuerte). Assessment Method.	88
Mapa 18. Clúster 4 (amarillo fuerte). Evaluation.	89
Mapa 19. Clúster 5 (morado). Participatory Evaluation.	90
Mapa 20. Clúster 6 (azul cian). Program Evaluation.	91
Mapa 21. Clúster 7 (anaranjado). Evaluation Capacity Building.	92
Mapa 22. Clúster 8 (marrón). Methodology.	93
Mapa 23. Clúster 9 (rosa). Planning Method.	94
Mapa 24. Clúster 10 (salmón). Participatory Research.	95
Mapa 25. Clúster 11 (verde pistache). Theory Based Evaluation.	96
Mapa 26. Clúster 12 (gris). Formative Evaluation.	97
Mapa 27. Clúster 13 (amarillo claro). Evaluation Studies as Topic.	98
Mapa 28. Clúster 14(rosa claro). Theory Based Evaluation.	99
Mapa 29. Clúster 15 (azul cielo). Performance Assessment.	100
Mapa 30. Clúster 16 (anaranjado claro). Program Development.	101
Mapa 31. Clúster 17 (marrón). Evaluation Use.	102
Mapa 32. Clúster 18 (rosa). Outcome Assessment.	103
Mapa 33. Clúster 19 (gris). Best Practices.	104

RESUMEN

Hoy en día, se tienen establecidos cuatro enfoques de evaluación: experimental, costo-beneficio, análisis de sistemas y cualitativo. Sin embargo, estos enfoques han ido evolucionando a lo largo de tiempo. Es por esto que, mediante un Estado del Arte del Campo Disciplinario de la Evaluación, se han identificado aquellas líneas de investigación o vacíos y necesidades que a lo largo de 10 años no se han desarrollado sobre evaluación, y que hasta el 2019 no se cuenta con información suficiente o documentada del seguimiento de éstas. Además, se encontró que los enfoques predominantes de evaluación son el cualitativo, cuantitativo y experimental.

Palabras claves: *evaluación, estado del arte, líneas de investigación, planeación, sistemas, educación, programas sociales, enfoques.*

ABSTRACT

Today, four evaluation approaches are established: experimental, cost-benefit, systems analysis, and qualitative. However, these approaches have evolved over time. This is why, through a State of the Art of the Disciplinary Field of Evaluation, those lines of research or gaps and needs that have not been developed on evaluation over 10 years have been identified, and that until 2019 are not counted with sufficient or documented information on their monitoring. Furthermore, it was found that the predominant evaluation approaches are qualitative, quantitative and experimental.

Keywords: *evaluation, state of the art, lines of research, planning, systems, education, social programs, approaches.*

INTRODUCCIÓN

Esta investigación tiene como principal objetivo estudiar la evolución de los enfoques, conceptos, metodologías y técnicas en cuanto al campo de la evaluación, y a partir de esto, identificar aquellas líneas de investigación que se han generado en un periodo de 10 años, para encontrar las necesidades en relación con la generación de conocimiento en este campo de estudio. Para cumplir con el propósito, el trabajo está dividido en seis capítulos y un anexo.

El primer capítulo de antecedentes muestra la conceptualización de la Evaluación desde el punto de vista de Stufflebeam, Madaus y Kellaghan con los siete periodos del desarrollo de la evaluación de programas. Del mismo modo, se presentan dos cuadros resumen de las generaciones y enfoques de la evaluación. El segundo capítulo establece los objetivos principales de esta investigación y lo que se espera obtener al final de este trabajo. En el tercer capítulo se describe la forma en que se establecieron las referencias conceptuales para realizar la búsqueda de información; además, se mencionan las principales fuentes de información como los repositorios digitales de donde se obtuvieron los artículos estudiados como su información bibliométrica. En el cuarto

capítulo se hace referencia al uso de la bibliometría para identificar grupos de documentos con patrones similares que sirven, a su vez, para analizar la estructura, el desarrollo y la evolución, en este caso, del campo de la evaluación, mediante el uso de los llamados mapas bibliométricos o de la ciencia.

Posteriormente, en el capítulo cinco se muestran los mapas de la ciencia construidos a partir de la información bibliométrica de los artículos, tanto para los *journals* de evaluación como para los de sistemas y planeación, y educación, y con éstos, representar la estructura de conocimiento en torno del campo de evaluación. Además, se menciona el uso de la herramienta VOSviewer como apoyo en la construcción de las redes bibliométricas. Finalmente, en el capítulo seis se presentan los comentarios y hallazgos producto de la revisión bibliográfica de las revistas especializadas. En el anexo A se presentan los mapas de coocurrencia de los clústeres sobre los *journals* de evaluación.

ANTECEDENTES

Una evaluación, es básicamente un juicio de valor. El objeto que se evalúa puede ser una propiedad tangible cuyo valor se cuenta en pesos y centavos; un ser humano cuya bondad es juzgada de acuerdo con sus defectos y virtudes; una actividad o programa cuyo éxito es medido por su desempeño, o incluso todo un sistema social o partes de ese sistema cuyo funcionamiento puede ser evaluado en términos de suavidad de operación o conflictos (Suchman, 1967).

Las personas evalúan y son evaluadas continuamente, como si se tratara de un proceso natural presente en muchos aspectos de la vida cotidiana, es decir, evalúan, valoran sin esfuerzo alguno y casi sin darse cuenta de ello. Aunque a diferentes niveles, también ocurre en los ámbitos profesionales. Por supuesto, la evaluación ha avanzado hacia una mayor amplitud teórica y sofisticación en las últimas dos décadas (Shadish, Cook, & Leviton, 1991).

En la vida social, la mayoría de las decisiones involucran más previsión, planeación y reflexión que en la evolución biológica. Se ha observado la evaluación social con la

selección de personal en China. También se tiene evidencia de las evaluaciones de los últimos 200 años (Shadish, Cook, & Leviton, 1991).

Nuestra preocupación es con las teorías, enfoques y prácticas de evaluación después del año 1960, especialmente en relación con la evaluación de políticas sociales. Estos esfuerzos tienen raíces intelectuales en trabajo anteriores, especialmente por Tyler (1935) en educación, Lewis (1948) en psicología social; y Lazarsfeld (Lazarsfeld & Rosenberg, 1955) en sociología. Ellos, también tienen raíces en el rápido crecimiento económico en Estados Unidos después de la Segunda Guerra Mundial, en el rol intervencionista que el gobierno federal de dicho país tomó en la política social durante la década de 1960, así mismo, en el creciente número de graduados en ciencias sociales que se interesaron en el análisis de políticas, la investigación de encuestas, los experimentos de campo y la etnografía. Todo esto preparó el escenario para la evaluación moderna de programas (Shadish, Cook, & Leviton, 1991).

La evaluación del programa a menudo se ve erróneamente como un fenómeno reciente. Forrester sitúa su primera manifestación histórica en el siglo II a. C, tratándose de una práctica china para seleccionar funcionarios; se trata de las primeras prácticas selectivas de evaluación oral. Lemus (citado en Alcaraz, 2015) añade que algo así como un cuestionario de evaluación que utilizaron Sócrates y otros maestros de la época en sus prácticas de enseñanza, alrededor del siglo V a. C. sitúa los sistemas de evaluación educativa entre los siglos V y XV d.C. No tan lejos, Lemus nombra a dos países precursores de la evaluación, durante el siglo XIX: Estados Unidos, en 1845, comienza a aplicar los test de rendimiento a estudiantes, con objeto de contribuir a la educación

de los estudiantes, y Gran Bretaña, donde existieron comisiones para evaluar los servicios públicos.

Muchas personas datan la evaluación de programas a principios de la década de 1960 con la infusión por parte del gobierno federal de grandes sumas de dinero en una amplia gama de programas de servicios humanos, incluida la educación. Sin embargo, la evaluación de programas tiene una historia interesante que precede en al menos 150 años a la explosión de la evaluación durante la era de la Gran Sociedad del presidente Johnson y el surgimiento de la evaluación como una profesión en proceso de maduración desde los años sesenta (Stufflebeam, Madaus, & Kellaghan, 2002).

Stufflebeam, Madaus y Kellaghan (2002) describen siete periodos del desarrollo de la evaluación de programas:

- i. **Edad de la Reforma (1800-1900).** Es la edad en la que se intenta reformar la educación y los programas sociales, y a las mismas agencias gubernamentales en Inglaterra y Estados Unidos, haciéndose intentos serios por realizar las evaluaciones. Las de las agencias se hacían gracias al empleo de métodos cualitativos y de naturaleza impresionista. El más empleado era el de acreditación que examinaba las evidencias y testimonios, y del juicio de un comité evaluador integrado por expertos, se emitían los resultados. En estos años, también se inició la práctica de enviar anualmente inspectores para supervisar los trabajos educativos. La evaluación era vista como una caja negra en la cual sólo interesaba valorar los productos. Este periodo se caracteriza por emplear el

método cualitativo de acreditación usando el juicio profesional de los expertos (Sánchez, 1994).

- ii. **Edad de la Eficiencia y la Experimentación (1900-1930).** Bajo la influencia de la administración científica, la evaluación de programas sociales ahora se orienta hacia la sistematización de la estandarización y principalmente a la medición de la eficiencia. Se pone de manifiesto una clara tendencia de entender la evaluación como control. Se inicia el empleo de un nuevo método, el de las encuestas. Para 1914 ya existían encuestas comprehensivas, llamadas así porque buscaban medir todas las facetas de la vida educativa. Así se inicia el auge de las encuestas. Sin embargo, en ocasiones los resultados de las encuestas y de los test son utilizados como propaganda, llegándose a distorsionar su interpretación. La evaluación de programas se hacía omitiendo su entorno. Se buscaba la eficiencia mediante la productividad al estilo de una empresa industrial; no se consideraban los aspectos conductuales de las personas. Estas deficiencias hacían que se siguieran realizando evaluaciones cualitativas con miras a ser comprehensivas (Sánchez, 1994).

- a. Asimismo, en su artículo Alcaraz (2015) menciona una época identificada por Stufflebeam y Shinkfield (2005), y Guba & Lincoln (1982,1987), los cuales la denominaron como época pretyleriana o la primera generación de la evaluación, respectivamente. Dicha etapa está comprendida entre el 2000 a. C y el año 1930. En esta generación de la medición, definida así por Guba & Yvonna (1989), el papel del evaluador se caracteriza por

aplicar los instrumentos de forma técnica, es decir, el evaluador era el técnico experto de la aplicación de instrumentos de evaluación para medir cualquier atributo o conjunto de ellos. Los test informaban algo sobre los alumnos pero nada de los programas de formación. Por su parte, para Stufflebeam y Shinkfield, las evaluaciones realizadas antes de los años 30 estaban centradas en valorar el resultado de los programas educativos en los alumnos a través de la aplicación de test, encuestas, acreditaciones y comparaciones experimentales.

- iii. **Edad Tyleriana (1930-1945).** Es reconocida así por Stufflebeam y Shinkfield, y recibe este nombre por Ralph Tyler, quien es considerado como el iniciador de la evaluación educativa; su visión racional de la evaluación es entendida también en términos de control, pero bajo un enfoque conductista. Tyler entiende la evaluación como la valoración del grado en que los objetivos deseados han sido realizados, de acuerdo con los resultados alcanzados, como parte de un programa institucional. Esto es, se mide qué tanta congruencia existe entre los objetivos marcados y los resultados alcanzados, gracias al empleo de test estandarizados o diseño de experimentos. Es una perspectiva de eficacia. En síntesis, este periodo se caracteriza por la aparición de un nuevo enfoque teórico y, conductista de la evaluación, que emplea la experimentación y los test estandarizados en ambientes controlados para medir la eficacia de los resultados (Sánchez, 1994).
 - a. A esta etapa tyleriana se le conoce también según Guba & Lincoln (1982,1989) como la segunda generación de la evaluación, la descriptiva,

que comprende hasta 1957, es decir, esta segunda generación se compuso por dos periodos, la hasta ahora edad tyleriana, según Stufflebeam y Skinfield (2005), y siguiendo a estos autores, lo que se llamó la Edad de la Ignorancia o Irresponsabilidad Social (Alcaraz, 2015).

- iv. **Edad de la Ignorancia (1946-1957).** La evaluación se orienta a ser instrumento justificador de las necesidades de expansión de Inglaterra y Estados Unidos. Ya no se busca medir la eficiencia de los programas, sino dar mayor impulso al análisis de las necesidades presupuestales que los programas requieren. Debido a la gran expansión, se tienen grandes avances en el perfeccionamiento de las técnicas de estandarización y el mejoramiento de los principios estadísticos del diseño experimental. Otro impacto favorable resulta casi al término de este periodo, cuando los evaluadores recapitan sobre el dispendio que se ha producido, orientado de nuevo a sus esfuerzos; pero ahora añadiendo el análisis costo – beneficio a la evaluación de programas. Este periodo se caracteriza por el crecimiento desordenado de la estandarización de test y del diseño experimental (Sánchez, 1994).
- v. **Edad del Desarrollo (1958-1972).** Durante esta edad tiene lugar una serie de acontecimientos coincidentes con gran parte de la Guerra Fría. Estados Unidos se encuentra en plena carrera armamentística con la URSS, de modo que ésta lanza en 1957 el primer satélite artificial de la historia, el Sputnik, lo que provoca una reacción de desencanto con la escuela pública en Estados Unidos. Además,

se produce un incremento de la presión por la rendición de cuentas que alcanza su máximo esplendor a finales de los años 60 con la aparición de un nuevo movimiento que se denominó la era de la accountability o rendición de cuentas.

- a. La expansión corporativa derivada de la posguerra impacta la evaluación de los programas sociales. La idea es expandirse como industria, pero racionalmente para no caer en el dispendio de los últimos años. Ante la urgente expansión, los métodos de evaluación se aplican a gran escala. Primero, el enfoque de Tyler es usado para ayudar a definir los objetivos para los nuevos currículos y más tarde para valorar el grado en que fueron realizados. Segundo, son creados nuevos test estandarizados a nivel nacional para que reflejen mejor los objetivos y contenidos de los currículos. Tercero, el juicio profesional de los expertos es mejorado y, por el auge de las encuestas, es usado para valorar los objetivos y checar periódicamente los compromisos adquiridos. Finalmente, muchos evaluadores se dan a la tarea de realizar experimentos de campo con los nuevos currículos. Años después, haciendo un balance de la expansión, el gobierno norteamericano, reconoce que se tuvieron muchos fracasos al llevar los métodos a una escala mayor de la que fueron diseñados. Los test fueron insensibles a las diferencias: entre escuelas y/o programas; de clases y razas, entre los valores individuales y sociales; y a la influencia que imprimía la administración que evaluaba, etc. Entonces el gobierno promueve la creación del Comité Nacional de Estudio de la Evaluación en 1971, que hace un llamado nacional para generar nuevos desarrollos

teóricos y métodos de evaluación, así como nuevos programas de enseñanza para la formación de los futuros evaluadores. Ante el llamado, surgen nuevas concepciones y las existentes se modifican. Así, de la experimentación se pasa a la cuasi-experimentación o de los test estandarizados a los test referenciados, pues se da una nueva apertura a los enfoques cualitativos. Se reconoce la necesidad de evaluar no sólo los objetivos sino también las entradas, las salidas, la capacidad del servicio y el proceso de implantación y la relación con el contexto, tal como es el caso de los métodos comprensivos CIPP de Stufflebeam en 1983, o el mejoramiento en las técnicas de estudio de caso de Stake en 1970, etc. Sin embargo, no se tiene una respuesta inmediata ante estas propuestas.

- b. Este periodo es reconocido por Guba & Lincoln como la tercera generación de la evaluación, la del juicio: el evaluador se convierte en un juez manteniendo las cualidades de las generaciones anteriores (Guba & Lincoln, 1989). Dentro de esta generación aparece la evaluación por toma de decisiones que se basa en la sociología de la organización, el análisis de sistemas, y la teoría y práctica de la administración. En otras palabras, la tarea principal del evaluador consiste en obtener y procesar sistemáticamente la información a fin de que pueda ser usada por quienes han de tomar decisiones.

- vi. **Edad de la Profesionalización (1973-1983).** Durante esta época, se presenta una proliferación de modelos de evaluación. Siguiendo la clasificación de Guba &

Lincoln (1982,1989) como la cuarta generación, la sensible; y, según Stufflebeam y Shinkfield , la época de la profesionalización (Alcaraz, 2015).

- a. Esta generación, que pudiera llamarse constructivista, se considera respondiente porque se da como respuesta a un proceso interactivo de negociación, donde participan todos los involucrados, con sus intereses y necesidades, y es constructivista porque define la metodología fundamental que se emplea, sin negar la posibilidad al método interpretativo o al hermenéutico. El evaluador, dentro de este paradigma, pierde su característica de juez, convirtiéndose en un investigador de procesos capaz de construir hechos que le permitan comprender el acto de aprender y por lo tanto de valorarlo (Muñoz, 2007). Como una respuesta ante la crisis, los evaluadores ya no son exclusivamente psicólogos, sociólogos o pedagogos, sino que se empieza a dar una incursión de otros profesionales al campo: administradores, maestros, ingenieros, economistas, etc. Los métodos que continúan dominando en las instituciones son la cuasi-experimentación y el uso de test estandarizados. Cada vez aparecen nuevos enfoques: la evaluación por objetivos de Scriven en 1974, que no varía mucho de la visión de Tyler; la evaluación mediante equipos de debate de Stake y Gijerde en 1974 que busca la síntesis de los argumentos básicos de los programas; la evaluación responsiva de Stake en 1978 que es una síntesis de la visión de Scriven y de Stake y Gijerde; los experimentos de campo controlados de Stanley en 1972, que son estudios igualmente de campo y cuasi-experimentación; la

evaluación naturalística de Guba en 1978, basada en métodos de campo e intervención y acción, y muchas más como la evaluación efectiva de Guba y Lincoln en 1981; la evaluación por estudios de caso de Stake en 1978; la evaluación cualitativa de Patton en 1980; la autoevaluación con gestión de calidad de Kells, Maassen y Haan en 1992; la evaluación política de Marshall y Peters en 1981, etc. Asimismo, se modifican nuevamente los enfoques dominantes. Por ejemplo, Cook, en 1972, propone modificar la evaluación que emplea técnicas de investigación de operaciones por técnicas de análisis de sistemas. Otros enfoques de análisis de sistemas son el de Stufflebeam de 1972 con la teoría de decisiones, y los de Edward, Guttentag y Snapper en 1975, y Metfessel y Michael en 1973 con el análisis multi-atributo de utilidad y multi-criterio, etc. El resultado es la proliferación de enfoques de evaluación, en su mayoría técnicas para mejorar la evaluación y selección, descuidándose muchos aspectos conceptuales y de la práctica (Sánchez, 1994).

- vii. **Edad de la Expansión e Integración (1983-2001).** Este último periodo comenzó de manera algo negativa, ya que hubo una reducción considerable en el financiamiento de las evaluaciones durante los años del presidente Reagan. Muchas evaluaciones durante este periodo se orientaron hacia la reducción de costos y las cuestiones de costo-beneficio. Sin embargo, a medida que la economía creció, la evaluación como campo se expandió y se hizo considerablemente más integrada. La expansión se observa especialmente en el

desarrollo de sociedades de evaluación profesional en más de veinte países y en los próximos años junto con la comunicación y la colaboración de evaluadores de diversas disciplinas. En educación, el movimiento de reforma ha tenido un profundo efecto en la evaluación del programa. Las reformas propuestas, como las escuelas autónomas, los vales y la privatización de escuelas, todas basadas en la creencia de que introducir la competencia en el sistema conducirá a una mejora, actualmente se están evaluando (Miron, 1999; Horn y Miron, 1999; Peterson, 1998; Peterson, 1996). Tennessee realizó un verdadero experimento para reducir el tamaño de las clases y los resultados han influido en la política en varios estados. La rendición de cuentas y las evaluaciones de resultados se han convertido en algo común en Estados Unidos. Sin embargo, estos sistemas de responsabilidad deben evaluarse de forma independiente, ya que muchos establecen objetivos de mejora poco realistas, basados en el desempeño de las pruebas de los estudiantes. El movimiento de reforma basado en estándares ha llegado a 49 estados. Los marcos curriculares se desarrollan al igual que las pruebas para medir el progreso en el logro de los estándares contenidos en dichos marcos. Varios estados han vinculado el rendimiento de los estudiantes en estas pruebas a nivel estatal con las decisiones de graduación y retención. La Evaluation Research Society, que estaba compuesta principalmente por evaluadores de las ciencias sociales, se unió a la Red de Evaluación, compuesta principalmente por evaluadores educativos, para formar la nueva American Evaluation Association (AEA). El resultado ha sido una integración y fertilización cruzada de ideas y métodos de evaluación en todas las disciplinas.

Esto está bien descrito en el artículo de Scriven (1994a) sobre la evaluación como una transdisciplina. Además, ha habido una gran expansión del campo profesional de la evaluación. En 1995, la AEA centró su convención en la cooperación internacional en evaluación e invitó a los evaluadores de todo el mundo a asistir. La reunión fue un gran éxito y generó una participación cada vez mayor de los internacionales en las reuniones de AEA y otros trabajos. Además, se han establecido más de 20 asociaciones de evaluación en todo el mundo, con un aumento concomitante en revistas de evaluación provenientes de otros países (Stufflebeam, Madaus, & Kellaghan, 2002).

Estado del Arte del Campo Disciplinario de la Evaluación

Tabla 1–Generaciones de la evaluación

PRIMERA GENERACIÓN DE LA EVALUACIÓN (HASTA 1930)	SEGUNDA GENERACIÓN DE LA EVALUACIÓN (1930-1957)	TERCERA GENERACIÓN DE LA EVALUACIÓN (1957-1972)	CUARTA GENERACIÓN DE LA EVALUACIÓN (HASTA 1980)
Se asume el positivismo de las ciencias físico-naturales.	Tyler acuña el término de “evaluación educativa”.	La evaluación se asocia a la toma de decisiones.	Se produce una “eclosión” de modelos de evaluación.
Interés por la medición científica de las conductas humanas.	Se va más allá de la evaluación psicológica.	Interés por rendir cuentas (accountability). El alumnado sigue siendo sujeto directo de la evaluación pero, además, también el profesorado, los medios, los contenidos, las experiencias de aprendizaje, la organización, etc.	Entre esos modelos aparecen los cualitativos, impregnados por la lógica del constructivismo. El interés se basa en la comprensión y reconstrucción de los acontecimientos educativos.
A través de las aplicaciones de instrumentos de formación técnica (test, de inteligencia y personalidad).	A partir de la propuesta de B. Bloom se define el currículum en términos de conductas (a través de objetivos muy precisos). Predomina el uso de test.	Cronbach incluye los cuestionarios, las entrevistas y la observación como técnicas de evaluación.	Es necesario fomentar el intercambio de opiniones, valores y experiencias entre los participantes de un programa a través de la utilización de métodos participativos.
Valoración de resultados de los programas educativos en estudiantes a través de la aplicación de test, encuestas, acreditaciones y comparaciones experimentales.	La evaluación se encarga de verificar la consecución de los objetivos propuestos, con la finalidad de valorar la eficiencia de los programas, aunque la aportación de la evaluación a la mejora de la enseñanza es escasa.	La evaluación sirve para rendir cuentas, y éstas, deben tener incidencia directa en los programas, los proyectos o las instituciones. Las decisiones deben mejorar la enseñanza.	La evaluación es una herramienta que facilita el empoderamiento, la emancipación de los individuos.
Medición es sinónimo de evaluación, anteponiéndose el término medición.	El término evaluación se antepone al de medición. Se habla de evaluación y medición.	Se incluye el término “juicio”. Se pretende valorar el mérito o valor de los programas. El evaluador es un juez que emite juicios.	Deja de hablarse de “medición” para avanzar en el concepto de evaluación educativa.
Evaluación basada en la norma. Informa sobre el rendimiento del individuo en comparación con un grupo.	La evaluación es basada en criterios. Indica el rendimiento de un individuo en relación con un estándar.	Predomina la “evaluación criterial”. Scriven acuña los términos de “evaluación formativa” y “sumativa”, o “evaluación intrínseca” y “extrínseca”.	Aparecen modelos alternativos (evaluación respondiente de Stake; democrática de McDonald; iluminativa de Parlett y Hamilton; la evaluación como crítica artística de Eisner, entre otros).

Fuente: (Alcaraz, 2015)

Tabla 2–Enfoques de evaluación

	EXPERIMENTAL	COSTO-BENEFICIO	ANÁLISIS DE SISTEMAS	CUALITATIVOS
CONCEPCIÓN DE EVALUACIÓN	El objeto a evaluar se considera aprehensible sobre el cual se puede experimentar Campbell & Stanley (1963)	El objeto de evaluación es una inversión y como tal deben tratarse BIRF(1961)	El objeto de evaluación se considera como recurso que puede ser aprovechado de manera óptima para alcanzar los objetivos marcados IIASA (1952)	El objetivo de evaluación es inaprehensible e imposible de ser encuadrado en un determinado patrón Guba & Lincoln (1959)
METODOLOGÍA	<ul style="list-style-type: none"> • Enunciado del problema • Formulación de hipótesis • Definición de variables • Definición de formas de medición • Diseño de instrumentos • Control de variables • Selección de un diseño • Selección y asignación de sujetos a grupos • Correr pruebas y aplicaciones del instrumento • Tratamiento estadístico de los datos • Estructuración del reporte de investigación • Elaboración de inferencias • Interpolación • Generalización de los resultados • Reporte final 	<ul style="list-style-type: none"> • Definición de objetivos • Especificación de variables • Determinación del conjunto de precios • Diseño de indicadores • Aplicación de modelos económicos • Estimación de los costos y beneficios • Selección 	<ul style="list-style-type: none"> • Formulación del problema • Objetivos, restricciones • Diseño de atributos/criterios • Identificar alternativas • Valorar alternativas: costo-objetivos • Desarrollo del modelo • Toma de decisiones • Comparación y clasificación de alternativas • Selección • Implantación 	<ul style="list-style-type: none"> • Selección de equipo de evaluación • Identificar stakeholders • Debate entre stakeholders • Identificación de necesidades, demandas internas • Valorizaciones • Priorizar problemas no resueltos • Buscar documentación más sofisticada • Negociación • Reporte final
HERRAMIENTAS	Diseño de experimentos <ul style="list-style-type: none"> • Encuestas • Test • Cuestionarios 	Análisis económico <ul style="list-style-type: none"> • Análisis de los efectos • Precios de cuenta • Precios eficientes 	Teoría de decisiones y programación matemática <ul style="list-style-type: none"> • Árbol de decisiones • Saaty, Delphi • Electra 	Análisis histórico y político <ul style="list-style-type: none"> • Debates • Categorización • Encuestas • Dinámicas grupales

Fuente: (Sánchez, 1994)

OBJETIVOS

En este sentido, y al ver el interés de los investigadores que han tenido por los textos y estudios sobre la evaluación desde hace varias décadas, el objetivo de esta investigación es llevar a cabo un análisis cuantitativo y cualitativo del estado del arte del campo disciplinario de la evaluación de los últimos 10 años.

A partir del análisis bibliográfico se pretenden determinar los nuevos estudios técnicos, metodológicos y conceptuales, y líneas de investigación que se han publicado sobre la evaluación del 2010 al 2019. Este objetivo se concreta en los siguientes aspectos:

- Estudiar la evolución en el tiempo de los estudios, enfoques, conceptos, metodologías y técnicas actuales de Evaluación, es decir, la dirección que han tomado éstos a lo largo del tiempo.
- Establecer una relación de autores que se han citado en los documentos revisados en el periodo mencionado.
- Definir las líneas de investigación que se han generado en materia de evaluación en un periodo de 10 años, y abordar lo que falta por hacer en torno a esta área.

- Identificar vacíos o necesidades en relación a la generación de conocimiento en el tema que aborda esta investigación.
- Identificar qué estudio o enfoque predomina en las actividades de la evaluación.
- Generar un estudio que sirva como contribución para generar nuevos problemas o hipótesis de investigación, y así ofrecer un insumo para dar comienzo a cualquier investigación que de enfoques de evaluación se refiera.

BÚSQUEDA, REFERENCIAS CONCEPTUALES Y FUENTES DE INFORMACIÓN

Como se mencionó anteriormente, el presente estado del arte del campo disciplinario de la Evaluación, tiene como finalidad saber qué se ha publicado por otros investigadores y, sobre todo, tener un marco de referencia de los temas ya estudiados, y de aquellas tendencias de desarrollo sobre el tema de nuestro interés. Del mismo modo, establecer un estado del arte sobre las nuevas aportaciones conceptuales, metodologías y técnicas de evaluación como campo de conocimiento que se han publicado entre los años 2010 y 2019.

Este proceso llevó a buscar, básicamente, en sitios web las distintas fuentes de información, para así poder seleccionar de forma crítica la información más relevante y poderla administrar de forma eficaz. La metodología de esta investigación está basada en una revisión bibliográfica de publicaciones y el análisis de los datos obtenidos de éstas

en relación con el tema estudiado, además del uso de repositorios digitales para la obtención de material bibliográfico.

El procedimiento de búsqueda y selección de los artículos consistió en un proceso ordenado, en donde, como punto inicial, se establecieron conceptos, palabras y autores clave que se consideraron concerniente al tema de evaluación, aquellas con las que se pudieran seleccionar y considerar como relevantes los artículos.

Es importante mencionar que se tomaron como base aquellos autores que han hecho aportaciones técnicas, conceptuales y teórico-metodológicas al campo de la evaluación. Por lo tanto, se consideran importantes a Michael Scriven y Donald Campbell por sentar las bases de una evaluación más racional mediante un conocimiento causal, y por desarrollar el enfoque de evaluación cuasi-experimental. Por su parte, Michael Scriven describe la evaluación como una transdisciplina. Carol Weiss y Robert Stake buscan el uso instrumental y más global de la evaluación, así como un cambio incremental. Independientemente, Stake adaptó el método de estudio de caso para su uso en la evaluación (Stufflebeam, Madaus, & Kellaghan, 2002).

Así mismo, Stake distingue dos formas de encarar la labor de la evaluación: tan, una denominada evaluación basada en criterios y estándares, que está orientada a las mediciones; la otra, llamada evaluación comprensiva, enfocada en la experiencia. La evaluación basada en estándares se fundamenta en el análisis de variables descriptivas, mientras que la evaluación comprensiva o interpretativa, lo hace sobre el conocimiento experiencial y personal en un espacio y es en tiempos reales, y con personas también reales (Araujo, 2014).

Lee Cronbach y Peter Rossi, después de señalar una crisis en los paradigmas de la evaluación, buscan una integración entre ellos. Individualmente, Lee Cronbach trabajó por una evaluación más participativa e integral que disminuyera el impacto del reduccionismo. Del mismo modo, Peter Rossi y Howard Freeman buscan una evaluación comprehensiva y una metodología para la conducción de los problemas, el desarrollo de los objetivos, el diseño de los programas y su implantación, además, dicen que el enfoque del costo-beneficio a programas sociales es preferentemente adecuado a evaluaciones ex-ante y de carácter local y exclusivamente monetario (Sánchez, 1994).

Michael Patton se caracteriza por señalar que existen grandes debates entre los paradigmas de evaluación, fundamentalmente los generados entre las dos grandes tendencias: la cuantitativa (análisis experimental) y la cualitativa. También se considera importante por la utilización del pensamiento de sistemas para la evaluación de programas sociales, así como el empleo de escenarios como una técnica que puede desarrollar el área. Juntos, Patton, Guba Edgon y Lincoln Yvonna, buscan un enfoque más integral y participativo, no objetivista y operativo.

Dos contribuciones dignas de mención que han buscado ofrecer una respuesta sistemática a la evaluación son los modelos CIPP de Daniel Stufflebeam y UTOS de Cronbach. El modelo CIPP (Context, Input, Process, Product) busca una evaluación comprehensiva de los programas, al evaluar de modo ordenado primero el contexto, luego los insumos, posteriormente el proceso y por último el producto.

El modelo UTOS (Units, Treatments, Observations, Settings) de Cronbach sugiere que la manera de hacer una evaluación comprehensiva es buscando un diseño muestral

apropiado, la cual busca una validez interna al ser consistentes las unidades, los tratamientos, las observaciones y la contextualización de las situaciones (Sánchez, 1994).

Del mismo modo, Werner Ulrich desarrolló la idea de “sistema crítico” bajo un marco conocido como Heurística de Sistemas Críticos, el cual se puede usar para evaluar cuán bien una intervención ha manejado cuatro áreas de selección de los límites críticos: impacto, control, conocimiento y legitimidad.

La revisión bibliográfica que se presenta se realizó a partir de buscadores, que incluyen Google y la Biblioteca Digital UNAM. Ésta forma parte de la oferta de servicios a la comunidad universitaria, tiene a través de la Biblioteca Digital un acervo bibliográfico en línea que incluye un millón de recursos digitales, 110 bases de datos, integrado entre otros, por libros, artículos y revistas. Además, se tiene disponible una herramienta de búsqueda simultánea en diversos catálogos y bases de datos que, por medio de una caja de búsqueda inspirada en el modelo de Google, proporciona resultados inmediatos y ordenados por relevancia.

La Biblioteca Digital cuenta aproximadamente con 419,314 libros, 23,389 revistas, 139,121 artículos de revistas latinoamericanas de ciencia y tecnología, 129,594 artículos de revistas latinoamericanas de ciencias sociales y humanidades, 2,721 recursos de fondo antiguo, 274,762 tesis, 13,792 mapas, 545 disposiciones legales del siglo XIX, 530 folletería del siglo XIX-XX y 114 bases de datos.

Asimismo, al no encontrar todas las revistas en la base de datos de la Biblioteca Digital UNAM se introdujo el nombre de las revistas en el buscador de Google, lo cual

direccionó la búsqueda a varios sitios. Uno de ellos es SAGE Publishing, el cual es un proveedor internacional líder de contenido innovador y de alta calidad que publica más de 1000 revistas y 800 libros nuevos cada año, que abarca una amplia gama de áreas temáticas; contiene una creciente selección de productos de biblioteca que incluye archivos, datos, estudios de casos y videos.

Otro sitio en donde se encontraron revistas es Elsevier, que al igual que SAGE Publishing, es un proveedor de información biomédica que ayuda y acompaña a los profesionales en el desarrollo de la práctica diaria y la formación continua. No obstante, cuenta con publicaciones especializadas en las áreas de ciencias sociales y tecnología, tales como economía y finanzas, negocios, administración, contabilidad y marketing, y psicología.

De igual forma, la búsqueda de las revistas remitió al sitio web de ScienceDirect, el cual es la plataforma para consultar la base de datos de revistas y libros editados por Elsevier, considerado el mayor proveedor mundial de información científica, técnica y médica. Publica más de 2,000 revistas, así como libros y bases de datos secundarias. La colección contiene más de 9.5 millones de documentos en texto completo, de alta calidad científica en 2,500 revistas y más de 11,000 libros, que abarcan diferentes áreas temáticas, entre las cuales se destacan: medicina, biología, física, matemáticas, ingeniería civil, ingeniería química, ingeniería informática, medio ambiente, ciencia de los materiales, ciencias sociales, sociología, artes y humanidades, economía, agricultura. La base de datos anualmente incluye casi medio millón de nuevos documentos. Comprende millones de artículos en una gran cantidad de archivos de información histórica antes de 1995.

Springer Books proporciona en línea, acceso al texto completo de títulos de revistas de Springer, así como revistas de otras editoriales. Los temas incluyen: arquitectura y diseño, ciencias de la conducta, biomédica y ciencias de la vida, negocios y economía, química y ciencia de los materiales, ciencias de la computación, la tierra y ciencias ambientales, ingeniería, humanidades, ciencias sociales y derecho, matemáticas y estadística, medicina, física y astronomía, informática profesional y diseño web.

Wiley Online Library proporciona acceso a más de 1,400 títulos de revistas en formato electrónico desde 1997 a la fecha. Cubre temas como agricultura, arquitectura, arte, química, computación, ciencias de la tierra, derecho, psicología, comportamiento humano, física, enfermería, medicina, etc.

Dialnet es una base de datos de producción científica hispana, creada por la Universidad de La Rioja, que integra múltiples recursos publicados en cualquier lengua que traten sobre temas hispánicos. Los contenidos incluyen artículos de revistas, capítulos de monografías colectivas, tesis doctorales, libros, etc. Su amplia cobertura convierte a Dialnet en la mayor base de datos de artículos científicos hispanos accesible de modo gratuito en internet.

Finalmente, Scopus, es una base de datos especializada en información científica, técnica, de ciencias médicas y sociales, que incluye las artes y las humanidades. Scopus abarca 29 millones de resúmenes de más de 18,500 revistas científicas (incluidas las 1,800 revistas de acceso abierto), 425 publicaciones comerciales, 325 series de libros, 250 actas de congresos y, 46 millones de registros. Los datos bibliográficos proporcionan información sobre los puntos clave de la base, tales como: identificador de autor;

Ceitation Tracker para un mejor control y seguimiento de citas en tiempo real; alertas, RSS feeds y HTML para mantenerse al día. Presenta interoperabilidad con SciVerse ScienceDirect, Reaxis y ProQuest CSA Illumina; y permite exportar los datos obtenidos mediante RefWorks, EndNote y BibTeX. Scopus es una herramienta que ofrece a los investigadores un recurso rápido, sencillo y completo para apoyar sus necesidades de investigación en cualquier ámbito.

El estado del arte del campo disciplinario de la evaluación se llevó a cabo a partir de la revisión bibliográfica de 20 revistas académicas especializadas en la temática de Evaluación, Sistemas, Planeación, Investigación de Operaciones y Educación (correspondientes a los años 2010-2019). Por otro lado, se seleccionaron estas revistas debido a que en la mayoría de éstas publican y/o editan los autores de referencia.

Cabe mencionar que los textos incluyen muchas otras referencias importantes, lo cual condujo a localizar fuentes adicionales, esto con la finalidad de reconocer a los autores en los que se basaron al fundamentar las ideas y argumentos de los artículos. Una vez que se revisaron todas las referencias pertinentes, se organizó, analizó e integró el material bibliográfico.

Las revistas especializadas consultadas fueron:

Evaluación:

- *American Journal of Evaluation*
- *Educational Assessment, Evaluation and Accountability*
- *Canadian Journal of Program Evaluation*
- *Evaluation Review*
- *Evaluation and Program Planning*
- *New Direction for Evaluation*
- *Educational Evaluation and Policy Analysis*

Sistemas y planeación:

- *Futures*
- *IEEE Transactions on Systems*
- *International Journal of General Systems*
- *Management Science*
- *OMEGA*
- *Technological Forecasting and Social Change*
- *Human Relations*
- *Interfaces*
- *Long Range Planning*
- *Policy Sciences*
- *Behavioral Science*

Revistas de educación:

- *Revista de la Educación Superior*
- *Revista de Investigación Educativa*

Es importante mencionar que debido a que los objetivos de este trabajo no es hacer una guía de búsqueda en bibliotecas digitales ni ahondar en temas bibliométricos; sólo se hará una descripción breve del procedimiento que permitió obtener información importante y relevante para cumplir los objetivos de esta investigación.

La búsqueda inició con la introducción del nombre de cada uno de los journals en el buscador de la Biblioteca Digital UNAM, la cual, en la mayoría de los casos, direccionó a la base de datos Scopus. En caso contrario, se visitaron otras bases de datos hasta lograr obtener la información bibliométrica de todas las revistas.

Scopus es una base de datos que permite la búsqueda, delimitación y exportación de archivos con formato csv. Es decir, en esta base de datos se encontraron la mayoría de los títulos de las revistas especializadas, así como las publicaciones hechas en el periodo que concierne a este trabajo (2010-2019). Además, Scopus permite exportar la información bibliométrica de todas y cada una de las publicaciones, tales como: año de publicación, título, autores, palabras clave, DOI, ISSUE, abstract, referencias, etc. Posteriormente, este archivo con toda la información necesaria, se introdujo a un software que permite visualizar mapas bibliométricos. Más adelante se explicará con mayor detalle.

En la siguiente tabla se muestra tanto el número de revistas como el de número de artículos obtenidos de las bases de datos consultadas.

Tabla 3 – Número de revistas por campo de conocimiento.

Tema de Revista	Número de revistas	Número de artículos encontrados
Evaluación	7	1006
Sistemas y Planeación	11	2938
Educación (español)	2	316
Total	20	3560

Fuente: Elaboración propia a partir de datos de diferentes bases de datos
(fecha de consulta 12/02/2020)

ANÁLISIS DE LA INFORMACIÓN

Para la identificación de las líneas de investigación y sus interconexiones, como se mencionó anteriormente, fue necesario realizar búsquedas en bases de datos, principalmente para comprender el estatus del campo de la evaluación, a partir del análisis de artículos de divulgación científica publicados en diferentes journals especializados.

En relación con lo anterior, una de las herramientas que se utilizó para el desarrollo de este trabajo fue la bibliometría que permitió realizar el análisis cualitativo de la bibliografía revisada. El principal objetivo de usar la bibliometría en esta investigación es identificar, entre otros aspectos, grupos de documentos con patrones similares que sirven, a su vez, para analizar la estructura, el desarrollo y la evolución, en este caso, del campo de la Evaluación.

Para comprender el uso de los mapas de la ciencia y su importancia en el desarrollo de esta investigación, se hará una breve introducción acerca de los mapas bibliométricos.

Es importante mencionar que, de la bibliometría, se utilizaron los **mapas bibliométricos, mapas de la ciencia o cienciogramas** como una representación espacial de la evaluación, que permiten delimitar diferentes áreas de investigación, con la finalidad de analizar la evolución de dicho campo. A partir de estos mapas, se puede comprender la representación gráfica de las diferentes unidades de análisis y sus relaciones (Gálvez, 2016).

Los mapas bibliométricos son una poderosa herramienta para el estudio de la estructura y dinamismo de un campo científico. Por ello, los investigadores pueden utilizarlos para obtener un mejor entendimiento del campo en el que están trabajando. Se distinguen varios tipos de mapas bibliométricos, cada uno visualiza la estructura de un campo científico desde diferentes puntos de vista. Algunos mapas, por ejemplo, muestran la relación entre autores o revistas basados en la cocitación, mientras que otros mapas muestran la relación entre palabras o palabras claves basadas en la co-ocurrencia.

Los mapas de la ciencia son la representación gráfica en dos o tres dimensiones de un campo de la ciencia. Ofrecen un paisaje o panorama de ésta, donde los elementos del mapa se refieren a temas o términos en el campo asignado. En estos mapas, los elementos se colocan relacionándose unos con otros, de tal manera que aquellos temas cognitivos que están relacionados entre sí, están situados en las proximidades, y los que no están relacionados o apenas lo están, se encuentran distantes (Guzmán Sánchez & Trujillo Cancino, 2013).

Para la generación de los mapas de la ciencia, se aplica un tipo de indicadores bibliométricos denominados indicadores relacionales. Esto es, la mayoría de los mapas de la ciencia se construyen básicamente por el principio de las relaciones de co-ocurrencia, o aparición conjunta, de dos unidades de información en un documento. Así, cuanto más cerca se encuentren dos elementos relacionados entre sí, más cerca se localizarán en el mapa.

En este mismo sentido, a partir de las relaciones de co-ocurrencia entre las unidades de información respectivas, se pueden realizar los siguientes análisis (Gálvez, 2016):

- a) Análisis de cocitación mediante el cual se mide la similitud entre documentos basada en las referencias bibliográficas de éstos. Se dirige, fundamentalmente, al estudio de los aspectos de la estructura intelectual o base de conocimiento de un área de investigación.
- b) Análisis de coautoría, por medio del cual se mide la similitud documental basada en las relaciones de coautoría de colaboración institucional, se dirige al estudio de los aspectos sociales y la dimensión internacional de un campo científico.
- c) Análisis de copalabras, por medio del cual se mide la similitud documental mediante el análisis de palabras. Básicamente, este método consiste en detectar palabras (que representan conceptos de los documentos en un dominio científico) para relacionar los documentos en los que aparecen según su grado de coocurrencia.

Estos análisis son lo que se dirigen al estudio de la estructura temática, semántica, conceptual y cognitiva de un área de investigación.

A continuación, se describen los diferentes análisis bibliométricos que se desarrollaron para esta investigación. Cabe mencionar que dichos análisis se hicieron por cada una de las áreas de conocimiento, es decir, un análisis bibliométrico tanto para las revistas de evaluación como para sistemas y planeación, y educación.

VISUALIZACIÓN Y DISCUSIÓN

Para realizar la visualización se utilizaron mapas bibliométricos en los que la representación de las relaciones entre las unidades de análisis se basa en la posición o distancia entre los nodos. La información bibliográfica obtenida de las diferentes bases de datos, se procesó con el software *VOSviewer*.

VOSviewer es un software desarrollado en el Centro de Estudios Científicos y Tecnológicos (CWTS) de la Universidad de Leiden (Holanda), y está diseñado principalmente para ser utilizado en el análisis de redes bibliométricas. Además, puede servir para crear mapas de publicaciones, autores o revistas sobre la base de una red de cocitación, o para crear mapas de palabras clave basadas en una red de co-ocurrencia; también, ofrece un visualizador que permite examinarlos, ofreciendo la posibilidad de exhibirlos de diferentes modos, cada uno enfatizando algún aspecto (Guzmán Sánchez & Trujillo Cancino, 2013).

El uso del software permite encontrar patrones presentes en los artículos analizados a través del uso de datos bibliográficos o datos de texto en diferentes formatos, los cuales una vez cargados en el software posibilitan, entre otros, los siguientes tipos de análisis:

- Coautoría por autores
- Coautoría por organizaciones (universidades, centros de investigación, facultades, escuelas, etc.)
- Coautoría por países
- Análisis de coocurrencia: palabras clave, palabras clave de autor, palabras clave del índice

Para esta investigación, se desarrollaron los mapas bibliométricos mediante el uso de VOSviewer y, particularmente, se elaboraron mapas basados en el análisis de cocitación y coocurrencia.

Cabe señalar que el *software* ofrece tres diferentes visualizaciones: red, densidad y superpuesta. Para esta investigación se utilizaron todas con el fin de mostrar la función de cada una de ellas. Por esta razón, no necesariamente se muestran las mismas visualizaciones para el análisis de los tres grupos disciplinarios (evaluación, sistemas y educación).

El análisis y el diagnóstico de la generación de conocimiento, en diversos temas de evaluación, son valiosos para conocer qué se ha investigado, qué se está estudiando e identificar tendencias que orientan futuras investigaciones. Para responder estas preguntas, se realizó el estudio bibliométrico de las investigaciones que se encuentran en diferentes bases de datos alrededor del campo de la evaluación. Al finalizar la búsqueda

de documentos, se alcanzó la recopilación de 3,560 artículos publicados entre el 2010 y el 2019. Los resultados obtenidos en las visualizaciones nos permiten comprender la estructura temática de las principales líneas de investigación en evaluación, y con base en las diferentes visualizaciones de las tres áreas temáticas en estudio, se pueden obtener grandes resultados.

Evaluación

Como primer lugar, se realizó la visualización del mapa bibliométricos del análisis de cocitación para los cuatro journals de evaluación, *Mapa 1*, que permite visualizar la estructura intelectual o base del conocimiento de este campo. Para esto, se decidió utilizar la visualización por densidad que ofrece el *software*. En este mapa se representan en color amarillo los autores que son citados con mayor frecuencia en los documentos y, en verde y azul, los menos citados.

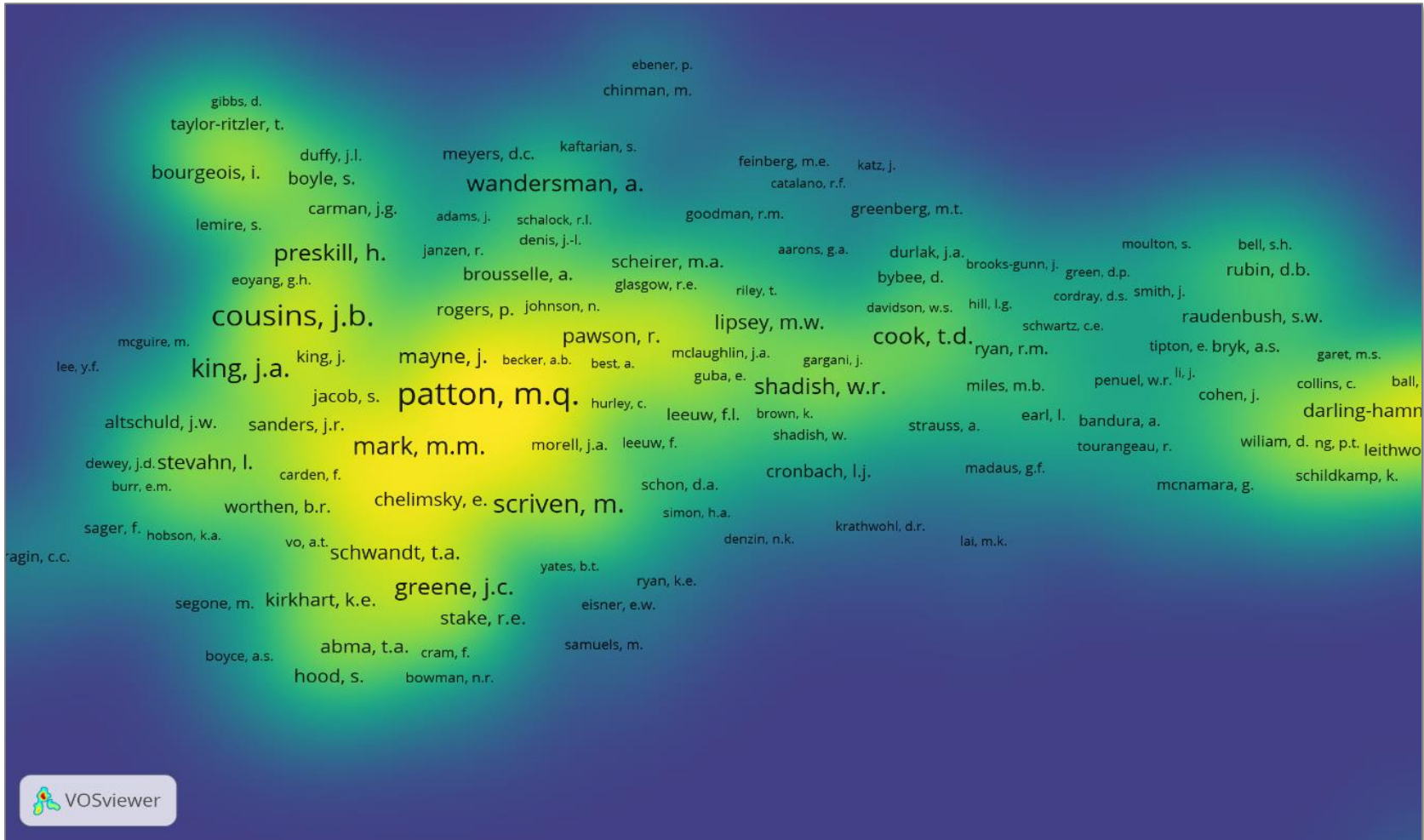
Con base en la visualización para este análisis, se muestran los autores más citados, independientemente de sus publicaciones. Se observa que el más citado es Michael Quinn Patton, consultor independiente con más de 40 años de experiencia en investigación aplicada y evaluaciones de programas; su trabajo se desarrolla en el ámbito de la investigación aplicada y evaluación en una amplia gama de temas, tales como educación, desarrollo de liderazgo, *developmental evaluation*, servicios humanos, medio ambiente, salud pública, efectividad de políticas, entre otros.

Del mismo modo, otros autores más citados en la generación de conocimiento del dominio de la Evaluación son J. Bradley Cousin, Jennifer C. Greene, Donald T. Campell, Thomas D. Cook, Carol Hirschon Weiss, William R. Shadish, Hallie Preskill, Abraham Wandersman, Lee J. Cronbach, Michael Scriven y Melvin M. Mark que, junto con Patton, son los mayores pilares en la estructura o base de conocimiento en el área de Evaluación en el periodo de tiempo determinado.

Adicionalmente, en el *Mapa 2* se utilizó la visualización de red para mostrar de mejor manera la relación existente entre la principal base del conocimiento y otros autores de las revistas de evaluación. El mapa de cocitación de autores mostró cinco comunidades de investigadores que se incluyeron dentro de la misma corriente de investigación (evaluación).

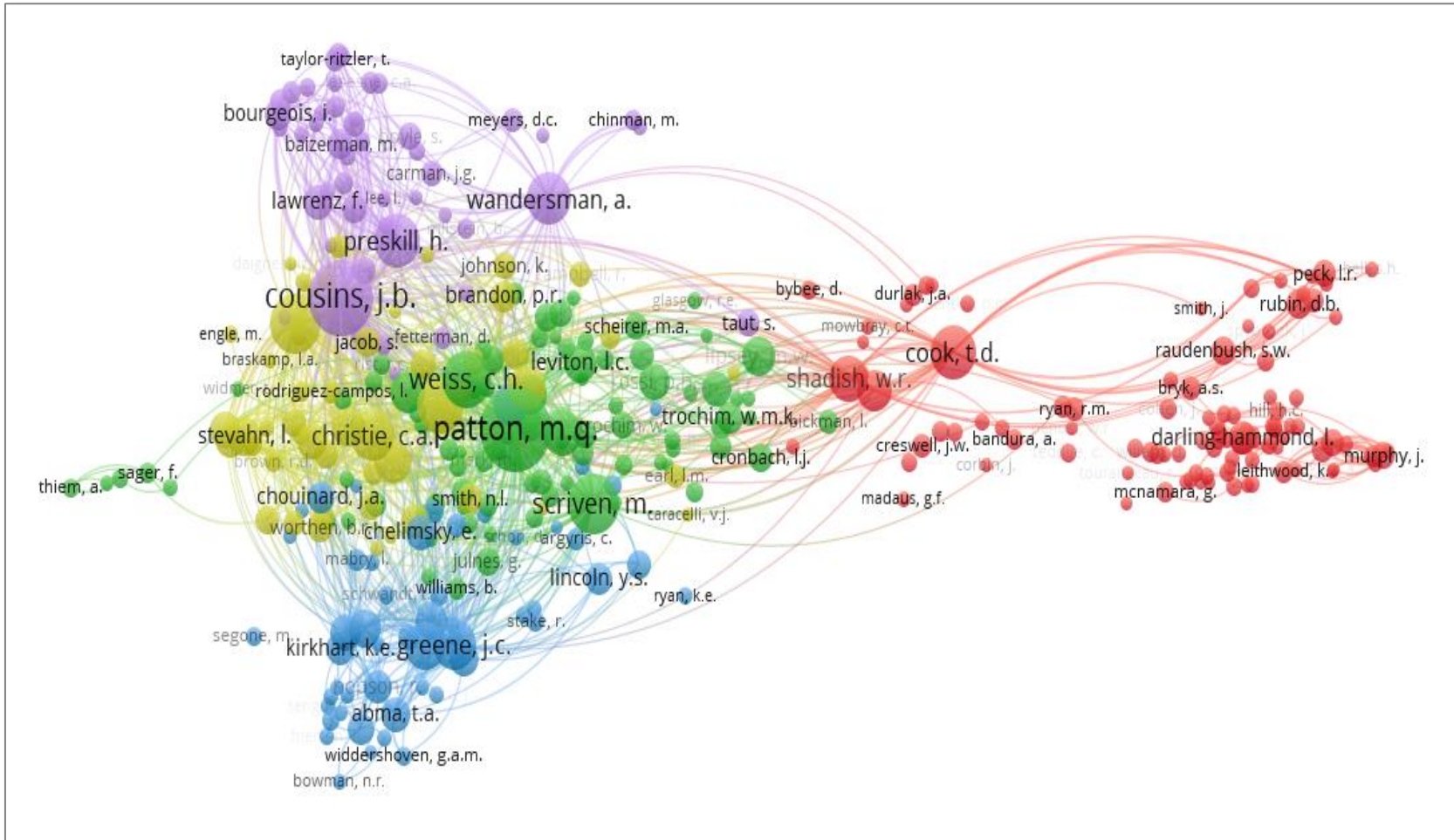
- **Clúster 1 (rojo).** Este grupo fue el que reunió un mayor número de autores: 103. Entre los más relevantes destacaron T. D. Cook, W. R. Shadish, D. T. Campbell y C. S. Reichardt.
- **Clúster 2 (verde).** Incluyó 90 autores, los más representativos son M. Q. Patton, S.I. Donaldson, M. Sriven, C. H. Weiss y R. Pawson.
- **Clúster 3 (azul).** Concentró 53 autores, los más influyentes fueron J. C. Greene, E. R. House, R. E. Stake, T. A. Schwandt.
- **Clúster 4 (amarillo).** Éste se conforma por 50 autores, de los cuales J. A. King, M. M. Mark, C. A. Christine, M. C. Alkin, C. H. Weiss y J. B. Cousins son los más relevantes.
- **Clúster 5 (morado).** Este grupo incluyó 48 elementos, los más influyentes son J. B. Cousins, H. Preskill, F. Lawrenz, A. Wandersman y M.Q Patton.

Mapa 1. Mapa bibliométrico de densidad de cocitación de autores de las revistas de evaluación.



Fuente: Elaborado a partir de información de las diferentes bases de datos, utilizando VOSviewer (fecha de consulta: 12/02/2020).

Mapa 2. Mapa de redes de cocitación de autores de las revistas de evaluación.



Fuente: Elaborado a partir de información de las diferentes bases de datos, utilizando VOSviewer (fecha de consulta: 12/02/2020).

Una vez que se hicieron los mapas de cocitación, se realizó la visualización de co-ocurrencia, esto significa hacer un mapa de red en donde se muestran las palabras claves o keywords representativas en los títulos, resúmenes y descriptores de los artículos obtenidos. En otras palabras, el *Mapa 3* de co-ocurrencia permite representar e identificar los términos que representan los conceptos contenidos en el texto de un documento. Con lo anterior, es posible identificar clústeres o líneas de investigación existentes en este campo. De modo que, al hacer este análisis, las diferentes agrupaciones (clústeres) situadas en el centro del mapa indican una alta interrelación de las keywords que lo conforman, mientras que los clústeres situados en los márgenes o bordes del mapa indican una menor interrelación de las keywords.

Los agrupamientos o clústeres, que a continuación se presentan, son mapas basados en la posición o distancia en que las relaciones se representan como puntos cercanos. Cada uno de los colores de las etiquetas representa el grupo temático (clúster) al que las keyword pertenecen, y el tamaño de las etiquetas refleja el peso de las aquéllas.

Es importante mencionar que mientras mayor sea el tamaño de las etiquetas, mayor será también la frecuencia de aparición en el flujo de información. También, cuanto menor sea la distancia entre dos términos, mayor será la relación (coocurrencias). En los diferentes clústeres puede haber palabras con mucha frecuencia de aparición y cerca o distante incluso de otros clústeres.

Con base en el mapa de co-ocurrencia de *keywords* de las revistas especializadas de evaluación, se identifica 19 clústeres temáticos que, se espera, definan las principales corrientes de investigación en materia de evaluación. Sin embargo, se definieron principalmente cinco grupos distribuidos por colores: rojo, azul, amarillo, verde y cian, que, de forma general, evidencian tendencias referentes a *developmental evaluation*, métodos de evaluación, evaluación como profesión, prácticas de evaluación y evaluación de programas, respectivamente. Más adelante se hará la descripción de cada uno de los agrupamientos (los mapas de los 19 clústeres, se pueden apreciar con mayor detalle en el Anexo A).

- **Clúster 1.** Developmental Evaluation. Esta línea de investigación se centró en las diferentes consideraciones que se deben tener presentes para llevar a cabo una evaluación, tales como: teoría del cambio, pensamiento crítico, planeación regional, intervención compleja, pensamiento evaluativo, planeación, sistemas de planeación, métodos de medición, evaluación de procesos y evaluación realista.
- **Clúster 2.** Éste se centró en la práctica de evaluación relacionada con la investigación sobre ésta. Agrupó los siguientes términos: relación entre teoría y práctica, políticas de evaluación, metaevaluación, métodos de investigación, programas comunitarios, evaluación multinivel, validación externa, inferencia causal, evaluación de impacto y evaluación educativa.

- **Clúster 3.** Aspectos relacionados con los diferentes métodos de evaluación. Otras temáticas presentes en este grupo son evaluability assessment, programa de mejora, métodos combinados, evaluación del riesgo, auto evaluación, evaluación escolar y evaluación del valor agregado.
- **Clúster 4.** Aspectos profesionales de la Evaluación. Se puede inferir de este clúster que hay tendencia de investigación hacia la profesionalización de la Evaluación, la carrera de Evaluador, participación, educación del evaluador, entrenamiento en evaluación, auto evaluación y, también en la precisión de la medición.
- **Clúster 5.** Investigaciones en evaluación participativa. Este frente reunió investigaciones sobre evaluación interna, rol del evaluador, evaluación intercultural, evaluación externa, cultura organizacional, capacidad de evaluación, rendición de cuentas, evaluación colaborativa, uso de la evaluación, capacidad de evaluación organizacional, autoevaluación y cambios en la organización.
- **Clúster 6.** Esta corriente se centró en la evaluación de programas e incluye los procedimientos de ésta, el rol y conocimientos del profesional, la participación de la comunidad, la evaluación y los programas basados en la teoría y evaluación de los problemas ambientales. Cabe mencionar que el nodo *Program Evaluation* es el que mayor peso tiene en todo el mapa; esto quiere decir que la frecuencia de aparición de este concepto en las investigaciones es alta.

- **Clúster 7.** Desarrollo de capacidades de evaluación. Esta línea de investigación se relacionó con la participación de los *stakeholders*, espacios de evaluación, investigación de la evaluación, evaluación de la enseñanza, nuevos evaluadores, evaluación de entrenamiento, investigación exploratoria y evaluaciones de impacto.
- **Clúster 8.** Esta línea se centró en el ítem Metodología, que tiene relacionadas las investigaciones enfocadas al marco conceptual, la evaluación de sistemas, *button-up approach*, la teoría de sistemas, el método de mapeo, el pensamiento de sistemas y el diseño de la evaluación.
- **Clúster 9.** Método de planeación es el ítem central de este clúster, que se relaciona con la investigación académica, la competencia y el entrenamiento del evaluador, el análisis de impacto, los instrumentos de validación, la necesidad de evaluación, la planeación de programas y los programas de investigación.
- **Clúster 10.** Este frente se centró en la investigación participativa que incluye los ítems de Métodos de evaluación, Participación, Uso de procesos, Sistemas, Simulación, Enfoque participativo, Fidelidad de implementación y Sistemas adaptativos complejos.
- **Clúster 11.** Este grupo temático se centró en la evaluación basada en la teoría, y los ítems con mayor peso fueron: Teoría de planeación, Solución de problemas,

Prácticas de planeación, Comunicación interdisciplinaria, Condiciones sociales y Moralidad.

- **Clúster 12.** En esta línea se situaron los estudios de evaluación formativa que se relacionan con los de educación para la salud, influencia de la evaluación, retroalimentación formativa, evaluación de precisión, credibilidad y metodología de la evaluación.
- **Clúster 13.** Este agrupamiento se centra en los estudios de evaluación e incluye ítems como Investigación colaborativa, Evaluación multisitio, Desarrollo metodológico, Evaluación de resultados, Teoría sociológica e Involucramiento de los *stakeholders*.
- **Clúster 14.** Este grupo se centró en investigaciones relacionadas con *stakeholders*, cuyos ítems más cercanos son: Evaluación de desarrollo, Administración y planeación del entorno, Instrumentos de evaluación, Evaluación y monitoreo, y Modelos estadísticos.
- **Clúster 15.** Éste se centra en el ítem Evaluación de desempeño, cuyos ítems más cercanos son: Análisis de datos y Cambio de sistemas; mientras que los más alejados son Políticas de evaluación, Estadísticas y Cooperación internacional.

- **Clúster 16.** Esta línea de investigación se centró en el desarrollo de programas con temas relacionados como evaluaciones multisitio, evaluación de proyectos y marco de evaluación.
- **Clúster 17.** Este pequeño clúster se centra en el uso político de la evaluación, principalmente en la formulación y participación política.
- **Clúster 18.** Este frente se centra en las investigaciones relacionadas con la evaluación de resultados, así como la evaluación de la investigación, la administración de proyectos y las evaluaciones de programas.
- **Clúster 19 (gris).** Finalmente, este frente de investigación se vincula únicamente con las mejores prácticas de evaluación.

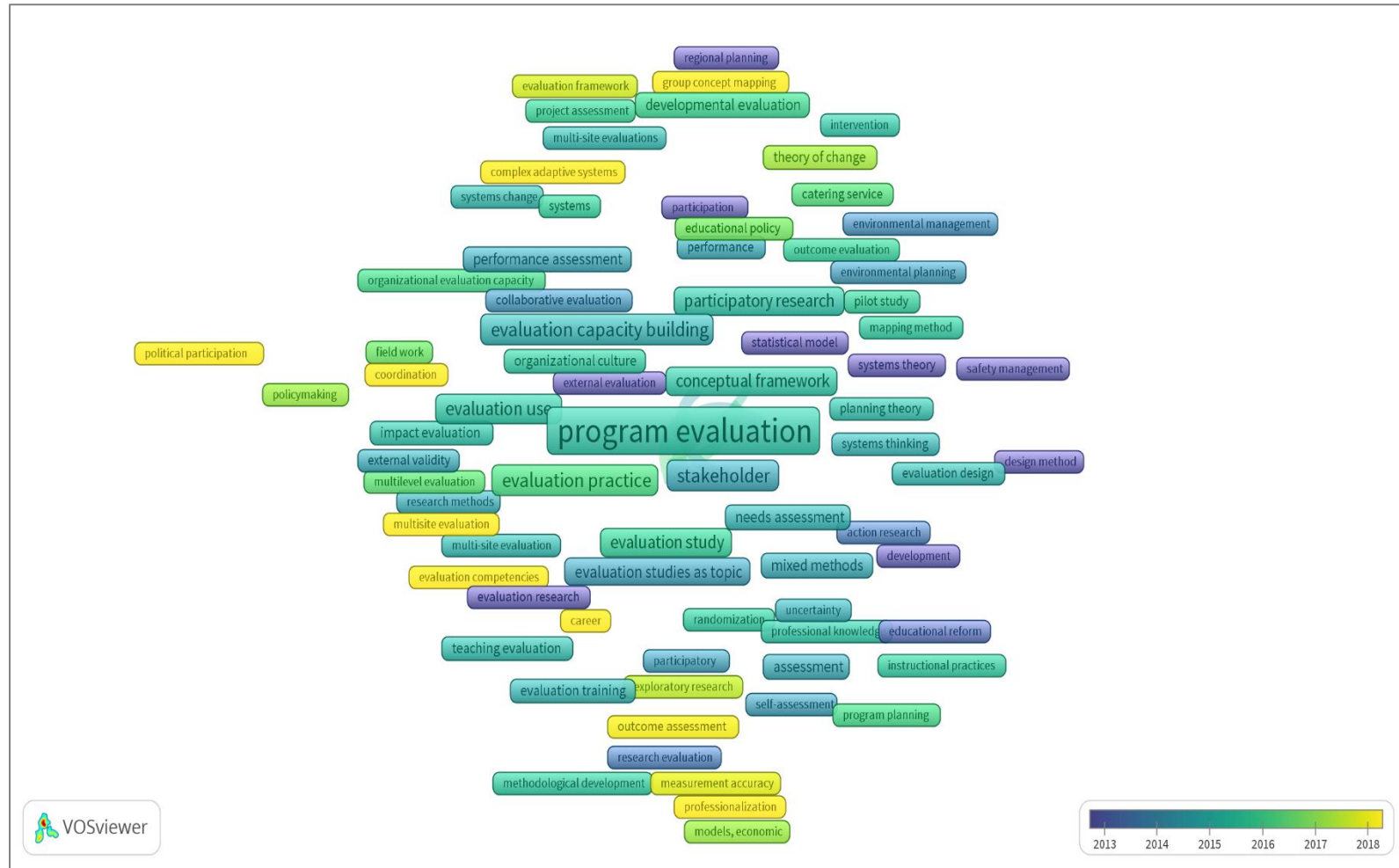
Posteriormente, para identificar los nuevos términos o ítems dentro del lapso en estudio, Vosviewer ofrece la visualización superpuesta u *Overlay Visualization*, en la cual se aprecia la evolución histórica que han presentado las investigaciones y, por medio de una escala de colores, se muestran tanto los términos más antiguos como los más recientes. En el *Mapa 4* se muestra la visualización de los términos y el año en el que fueron citados: en color azul los conceptos anteriores al 2013, y en amarillo los más recientes.

De lo anterior, se puede observar que las investigaciones de los últimos 10 años se han enfocado, inicialmente, en la planeación regional con un enfoque participativo, así como en la evaluación externa, incluyendo modelos estadísticos, teoría de sistemas, métodos

de diseño, investigación de la evaluación y evaluación de la investigación. Posteriormente, el tema de investigación más frecuente fue el de la evaluación de programas, nodo principal del mapa, el cual estuvo relacionado principalmente con tendencias como la inclusión de los *stakeholders*, pensamiento de sistemas, teoría de planeación, cultura organizacional, necesidad de evaluar, investigación participativa, evaluación de impacto, cambio de sistemas, diseño de evaluación. Después, los investigadores tornaron su interés hacia las investigaciones relacionadas con la práctica de evaluación, al mismo tiempo que las temáticas como la evaluación a niveles múltiples, planeación de programas, evaluación de proyectos, *developmental evaluation* y capacidad de evaluación organizacional.

Por último, se puede observar que las investigaciones más recientes se han orientado hacia conceptos o temas como la política educacional, teoría del cambio, evaluación multinivel, formulación de políticas, investigación exploratoria, marco de evaluación, modelos económicos, así como a la profesionalización de los evaluadores, participación política, coordinación, evaluación multinivel; y de nueva cuenta, competencias de evaluación, evaluación de resultados, sistemas adaptativos complejos y precisión de la medición.

Mapa 4. Mapa de redes de visualización superpuesta por año de publicación en el área de evaluación.



Fuente: Elaborado a partir de información de las diferentes bases de datos, utilizando VOSviewer (fecha de consulta: 22/10/2019).

En lo que respecta a la red de trabajo de los autores que han publicado en el área temática de evaluación, se obtuvieron los resultados en forma de nodos o grupos de trabajo de los autores; en total se identificaron ocho clústeres como se muestra en el *Mapa 5*.

En esta visualización, el tamaño de los círculos es proporcional al número de artículos publicados por el o los autores sobre evaluación. A continuación, se enlistan los agrupamientos encontrados.

- **Clúster 1 (rojo).** En este grupo, correspondiente a los autores M. C. Alkin, S. I. Donaldson, J. A. King, J. M. y F. Lawrenz, Alkin y King abordan el concepto de uso de la evaluación. Como campo, la evaluación siempre ha prestado atención al potencial de uso, tanto en la toma de decisiones como en el cambio de pensamiento de las personas. La amplia historia del campo tal como la conocemos hoy proviene de dos corrientes: una centrada en pruebas y mediciones, principalmente en educación, y otra centrada en métodos de investigación social, principalmente en relación con la utilización del conocimiento. El uso de la evaluación tuvo sus raíces en ambas corrientes, dando como resultado tres categorías amplias para discutir los usos de los resultados de la evaluación: instrumental, conceptual o iluminación y uso simbólico. La categoría adicional de uso del proceso, agregada años después, destacó la utilidad potencial de la participación de las personas en el proceso de evaluación (Alkin & King, 2016).

Asimismo, M. C. Alkin, A. T. Vo y C. A. Christie consideran que el acto de valorar en una evaluación puede percibirse de diferentes maneras, es decir, consideran las múltiples perspectivas teóricas que rigen el comportamiento de un

evaluador y presentan una tipología de los roles de valoración del evaluador. Dentro de esta tipología, describen tres formas en las que los juicios de valor son alcanzados típicamente: sólo los *stakeholders*, los *stakeholders* y los evaluadores en conjunto, o sólo los evaluadores. Esta heurística ofrece una comprensión más explícita de cómo ocurre la valoración en la evaluación (Alkin , Vo, & Christie, 2012).

Por su parte, S. I. Donaldson utiliza el concepto de evaluación de empoderamiento para describir el desarrollo de enfoques de participación de las partes interesadas para la evaluación. Además, considera que el respeto de la evaluación de empoderamiento por el conocimiento de la comunidad, y el compromiso con el derecho de las personas para construir su propia capacidad de evaluación han influido en la corriente principal de la evaluación, particularmente en relación con el desarrollo de la capacidad de evaluación. Las contribuciones más significativas de la evaluación de empoderamiento al campo han sido mejorar el uso de la evaluación y la utilización del conocimiento (Donaldson, 2017). Posteriormente, en su publicación más reciente, debate temas importantes de la profesionalización de la evaluación, incluidos los beneficios potenciales, los efectos secundarios negativos, los procesos, las competencias y los procedimientos. Los resultados de entrevistas y una encuesta en línea demuestran que los miembros de la AEA consideran que los beneficios potenciales de la profesionalización son la confianza de los interesados y la reputación e identidad del evaluador; mientras que las preocupaciones sobre un posible efecto secundario negativo conocido como el "efecto limitante" (es decir,

algunos evaluadores considerarán ser enajenados en función de sus antecedentes, competencias, etc.) fueron expresadas por los participantes. Estos hallazgos recientes pueden informar la discusión en curso sobre la profesionalización y sugerir nuevas direcciones para futuras investigaciones sobre evaluación (Donaldson, 2019).

En 2011 F. Lawrenz, J. A. King y A. Ooms promovieron el uso y la influencia de la evaluación multisitio. A partir del estudio realizado por ellos, se encontró que las personas que participaban en las evaluaciones multisitio percibían su participación de manera diferente a la visión más tradicional de la participación en evaluaciones locales. El estudio también destacó la importancia de atender los siguientes temas: la interfaz entre el financiador y el personal local del proyecto; los ciclos de vida del programa y su evaluación, así como los del proyecto local y su personal individual; la información sobre quién controla los proyectos; las muchas tensiones que pueden existir entre los objetivos del programa a gran escala y sus proyectos más pequeños; la calidad percibida de la evaluación y los evaluadores, y las posibilidades de construcción de la comunidad y la creación de redes entre proyectos (Lawrenz , King, & Ooms, 2011). Del mismo modo, en la red que lo conforma A. M. Gullickson, J. A. King, J. M. LaVelle y J. M. Clinton, se inclinan hacia la educación del evaluador como una iniciativa y proporcionan una base para su evaluación. En él, los autores utilizan el modelo de contexto, entrada, proceso y producto () de Stufflebeam para enmarcar una revisión de la bibliografía y la práctica. Su análisis sugiere que hay mucho trabajo por hacer para comprender las necesidades de educación del evaluador; delinear

los estándares de calidad tanto en la educación como en la práctica; identificar los insumos y procesos más efectivos para abordar esas necesidades, y documentar sus impactos finales. El documento ofrece un llamado a la acción para mejorar la calidad, la consistencia y la integridad de este importante trabajo (Gullickson , King , LaVelle, & Clinton , 2019).

- **Clúster 2 (verde).** En esta red, los autores A. S. Boyce, J. A. Chouinard, J. B. Cousins, P. Milley y B. Szinjarto han desarrollado en el periodo de estudio, temas de investigación como el seguimiento del impacto de la práctica de la Investigación sobre Evaluación (RoE, por sus siglas en inglés). Estos autores se han dado a la tarea de encontrar investigaciones y utilizar el conocimiento para desarrollar un enfoque conceptual para evaluar el impacto de RoE en la política y la práctica de la evaluación. Específicamente, desarrollan una teoría de cambio para RoE como parte de un enfoque de análisis de contribución a la solución de problemas.
- **Clúster 3 (azul).** Para el grupo de los autores C. A. Christie, L. A. Fierro, S. Lemire y S. B. Nielsen las diferentes líneas de investigación que se han desarrollado en 10 años comprenden temas como patrones de evaluación del curso para evaluadores (Christie, Quiñones & Fierro, 2014). Así mismo, se discuten las similitudes y puntos de partida entre la ciencia de la mejora y la evaluación, de acuerdo con el uso, la valoración y los métodos (Christie, C.A., Lemire, S., & Inkelas, M., 2017). También, se han interesado por la necesidad de

comprender mejor cómo medir la capacidad de evaluación y sus posibles resultados, y así investigar en qué medida las evaluaciones de la capacidad de evaluación y la práctica de evaluación son similares o diferentes para las personas que representan el mismo programa (Fierro & Christie, 2017). Igualmente, las investigaciones se han llevado hacia la importancia de considerar la medida en que las políticas son informadas por la teoría, dados los propósitos relacionados con las teorías y políticas de evaluación. Debido a que estas últimas han surgido con mayor frecuencia hace relativamente poco tiempo, el trabajo conceptual y empírico sobre la conexión teoría-política-práctica es limitado. El interés de esta línea de investigación es desarrollar la intersección entre la teoría, la política y la práctica de la evaluación (Christie & Lemire, 2019).

- **Clúster 4(amarillo)**. En esta red, los autores T. Archibald, C. L. S. Coryn, K. A. Hobson y A. T. Vo se encontraron líneas de investigación relacionadas con el pensamiento evaluativo por sus siglas en inglés). Es un tema cada vez más importante en el campo de la evaluación, particularmente entre las personas involucradas en el desarrollo de capacidades de evaluación, además de que se proponen varios principios rectores y estrategias específicas para la enseñanza de ET, que se basan directamente en la investigación sobre la enseñanza del pensamiento crítico (Buckley, Archibald, Hargraves, & Trochim W., 2015). Asimismo, el pensamiento evaluativo se ha utilizado en el desarrollo de capacidades de evaluación BCE por sus siglas en inglés), diseñado para ayudar a los profesionales del desarrollo comunitario a trabajar de manera más efectiva

con suposiciones en el proceso de planificación, monitoreo y evaluación del programa, que además involucra la creación y análisis de modelos de teoría de cambio (Archibald, Sharrock, Buckley, & Cook, 2016).

- **Clúster 5 (morado).** Para los autores T. Azzam y M. R. Jacobson se enfocan en el futuro de la investigación sobre evaluación, RoE, la cual es utilizable y relevante para los evaluadores en ejercicio, la recopilación sistemática y el intercambio de datos de prácticas de evaluación, el desarrollo activo de métodos y herramientas que miden de manera factible y rigurosa diferentes aspectos de la evaluación, el uso de nuevas tecnologías en RoE y la difusión mediante sistemas nuevos y existentes que se dirigen a otros investigadores y profesionales de la evaluación.
- **Clúster 6 (neutro).** E. R. House se encuentra en la rama de valoración del Árbol de la Teoría de la Evaluación, y se centra en la justicia como igualdad. Examina la descripción del programa Mujeres Afirmando la Maternidad y considera qué enfoques podrían conducir a una evaluación igualitaria que ayude al programa en sus esfuerzos por expandirse. En particular, House propone un equipo de evaluación colaborativo, una investigación de factores causales, incluido el encuadre racial, y la presentación de hallazgos a través de escenarios e historias convincentes (House E., 2019). Del mismo modo, en el libro *Evaluating with Validity* (House, 1980) amplió la concepción de validez del campo de evaluación al afirmar que las evaluaciones deberían ser verdaderas, coherentes y justas. Las

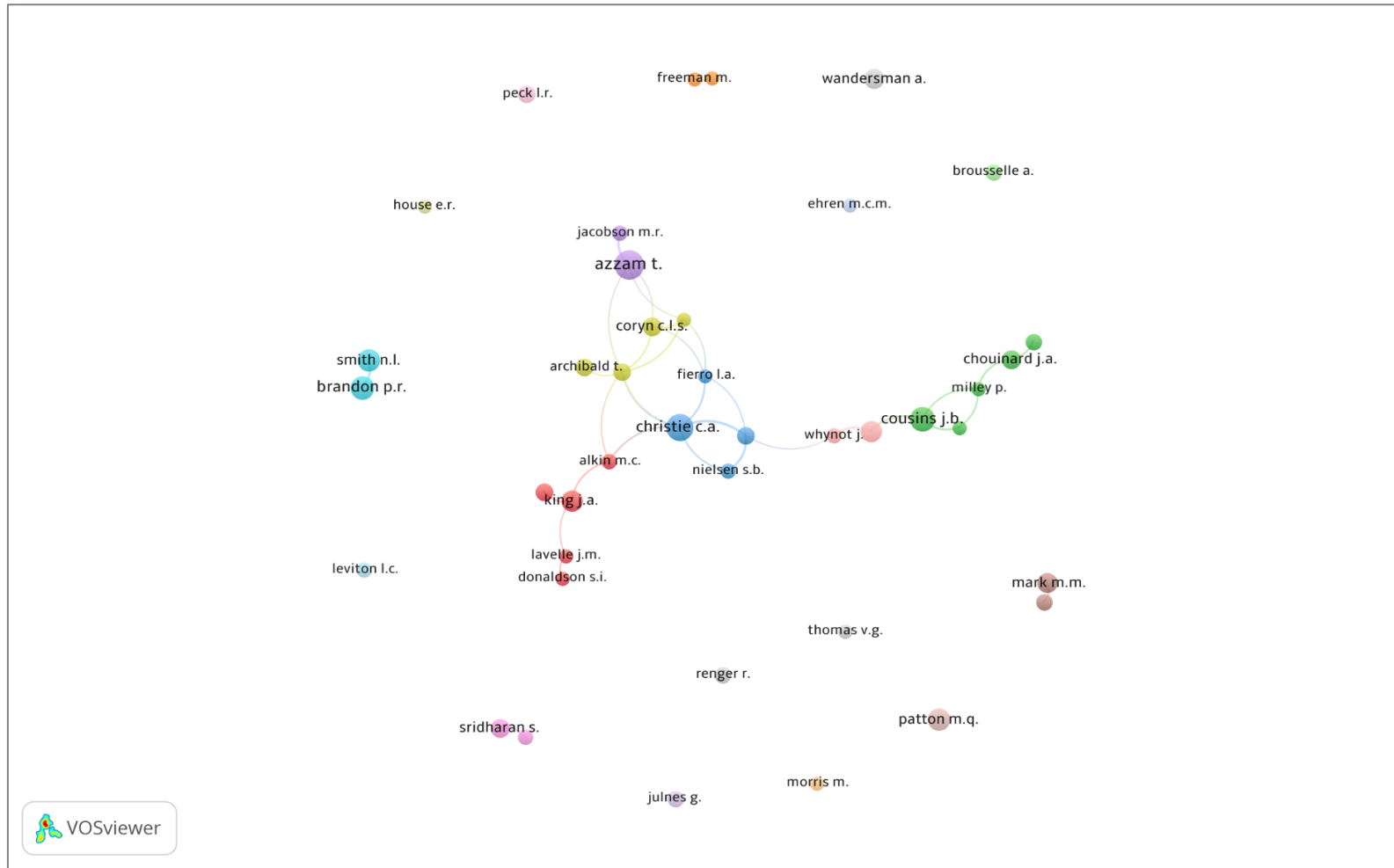
evaluaciones falsas, incoherentes e injustas no son válidas. Las ideas de trabajo fueron argumento, coherencia y política. Para el criterio de verdad, no fueron los diseños o las correlaciones los que determinaron la validez, sino más bien la de los argumentos que respaldaron el uso de los diseños y las correlaciones. La concepción más amplia de validez creció al abordar los problemas encontrados en la realización de evaluaciones. En su investigación, House traza los orígenes de las ideas y el contexto social del que surgieron. Sostiene que estos criterios aún se aplican, aunque el contenido de los criterios ha cambiado un poco y el contexto ha cambiado sustancialmente (House E., 2014).

- **Clúster 7 (neutro).** El autor G. Julnes centra su línea de investigación, desde el 2012 hasta el 2019, hacia el apoyo evaluativo para la valoración de programas y políticas que buscan servir al interés público. El autor aborda formas de administrar el uso de metodologías de valoración de acuerdo con los contextos en los que funcionan mejor. Esto, a su vez, respalda las conclusiones sobre los tipos deseados de equilibrio en métodos comunes de valoración en el campo de la evaluación (Julnes, *Managing valuation*, 2012). Así mismo, este autor promueve el profesionalismo y se centra en cómo los valores deben influir en la práctica de la evaluación y en cómo la comprensión de la valoración afecta la visión de la ética que debería guiar a los evaluadores (Julnes & Bustelo, 2017).
- **Clúster 8 (neutro).** M. Q. Patton y otros investigadores ayudan a los lectores a comprender y aplicar las técnicas de identificación y análisis de partes interesadas en el diseño de evaluaciones que mejoren el uso primario de los usuarios

principales; así mismo, presentan una gama de las técnicas más relevantes para identificar y analizar a los interesados en la evaluación (Bryson, Patton, & Bowman, 2011). De igual manera, Patton se enfoca en que la comprensión de los programas de mejora cuando los evaluadores entrenados, capacitados y observadores entran al campo, al mundo real donde se llevan a cabo dichos programas, al prestar atención a lo que sucede, documentar sistemáticamente lo que ven e informar sobre lo que aprenden. Menciona las razones para incorporar el trabajo de campo en diseños de evaluación. La investigación de Patton identifica debilidades sistemáticas en cómo se encargan y realizan las visitas al sitio, y propone estándares para mejorar la calidad y el uso (Patton M., *Evaluation in the Field: The Need for Site Visit Standards*, 2015). Continuando con lo anterior, en 2017 Patton junto con otros colegas, presentan qué ayudaría a los profesionales a aclarar su pensamiento sobre sus propias visitas al sitio y qué ayudaría en la planeación de la evaluación y la fase de diseño de su práctica (Chapman, Murphy, & Patton, 2017). Finalmente, basado en los trabajos anteriores, Patton presenta algunos estándares mínimos de control de calidad para las visitas al sitio, además de los principios de mejora también de la calidad, destinados a garantizar la excelencia. Esto como un llamado para hacer que el trabajo de campo de evaluación se base en estándares (Patton M., 2017). Para 2018, Patton reconoce la evaluación como ciencia, y con esto afirma que mejora la credibilidad y eficacia de ésta, además de que permite comunicar el papel del evaluador de manera más clara y creíble a quienes valoran la ciencia. También asevera que ver la evaluación como ciencia puede afectar la forma en que son

vistos, tratados y posicionados los evaluadores en la academia, el gobierno y los patrocinadores, y usuarios de la evaluación. Patton basa su análisis en las ramificaciones de la ciencia de la evaluación para el estado de ésta como profesión, rama de las ciencias sociales aplicadas, disciplina y transdisciplina (Patton M., 2018).

Mapa 5. Mapa de coautoría en el área de evaluación.



Fuente: Elaborado a partir de información de las diferentes bases de datos, utilizando VOSviewer (fecha de consulta: 12/02/2020).

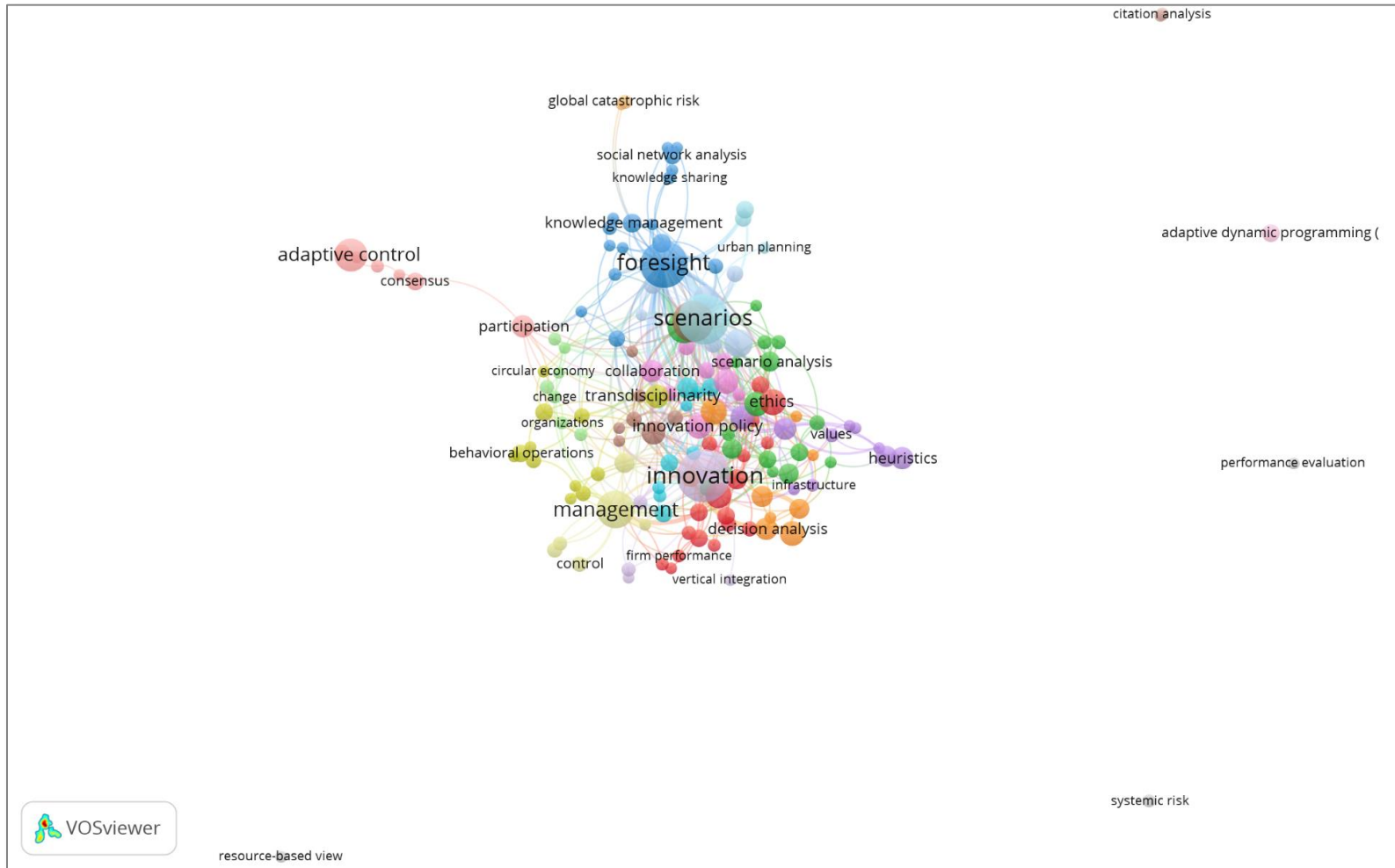
Sistemas

Para el caso de las visualizaciones de los 11 *journals* de sistemas y planeación, se llevó a cabo de una forma distinta. Es decir, para la construcción de los diversos mapas fue necesario excluir conceptos, ítems o palabras que no tuvieran relación con autores, definiciones y/o teorías de evaluación, o que fueran estudios de caso. Sin embargo, se identificaron aquellas relaciones existentes entre los ítems representativos en el campo de la evaluación, y con base a esto, se construyó el mapa correspondiente.

El objetivo de elaborar el mapa de redes para este campo temático es identificar los ítems relacionados con el o los conceptos más representativos sobre la evaluación, y así determinar las líneas de investigación alrededor de ésta con un enfoque de Sistemas y de Planeación.

En el *Mapa 6* se observa la red de cocurrencia de las palabras en los documentos encontrados en las diferentes revistas. De igual forma que en las visualizaciones anteriores, mientras más grande sea la etiqueta de las palabras, mayor es la frecuencia de aparición en el flujo de información. Para complementar la visualización del mapa anterior, en el *Mapa 7* se muestran en color amarillo aquellas palabras que se referencian con mayor frecuencia en los documentos y, en verde y azul se sitúan aquellas menos relevantes. Finalmente, en el *Mapa 8* se puede apreciar la evolución histórica que han presentado los documentos, mostrándose en azul los componentes del concepto anteriores a 2017 y en amarillo los más recientes.

Mapa 6. Mapa de redes de coocurrencia de palabras claves de las revistas de sistemas y planeación con aristas.



Fuente: Elaborado a partir de información de las diferentes bases de datos, utilizando VOSviewer (fecha de consulta: 12/02/2020).

Como se puede apreciar en los tres mapas anteriores, las etiquetas que sobresalen son las que contienen los conceptos de escenarios, innovación y pronósticos, los cuales por sí solos no dan mayor evidencia de ser líneas de investigación de nuestro interés. De igual manera, la visualización superpuesta de la red de coocurrencia tampoco muestra información relevante. Sin embargo, si hacemos el ejercicio de puntualizar el concepto *evaluation*, automáticamente arrojará la relación existente entre otros conceptos, y de esta forma se podrán hacer conjeturas que permitirán identificar las líneas de investigación alrededor de la evaluación en el campo de sistemas y planeación.

De acuerdo con lo anterior, los principales conceptos relacionados con evaluación (*evaluation*), que presentan mayor frecuencia de aparición en el flujo de información de estos campos de conocimiento son: escenarios, pronósticos, transdisciplinariedad, participación y *stakeholders*. No obstante, si organizamos estos mismos conceptos en orden cronológico, se aprecia que *stakeholder* es el ítem más reciente, seguido de participación, pronósticos, transdisciplinariedad, y escenarios como el más antiguo de los conceptos.

En el *Mapa 9* se muestran aquellos conceptos que presentan una relación con evaluación en una visualización superpuesta, es decir, además de mostrar la relación entre conceptos, se puede apreciar qué tan recientes son éstos.

Para el caso de las investigaciones que tratan el tema de las evaluaciones de nuevas tecnologías, éstas involucran técnicas participativas para la estructuración de problemas como la aplicación de entrevistas como medio de participación, en la cual se pueden aclarar problemas comunes y diversificados en función de la percepción de las partes interesadas clave e identificar otras adicionales.

En cuanto a las evaluaciones de la sostenibilidad ambiental de proyectos energéticos, la evaluación se lleva a cabo mediante un indicador de sostenibilidad ambiental que se puede utilizar en la comparación de escenarios, al mismo tiempo que se mide la efectividad de las alternativas propuestas. Al hacerlo, es posible utilizarlo como herramienta de toma de decisiones para los planeadores del sector energético. Así mismo, para estas evaluaciones de sostenibilidad ambiental, utilizan técnicas como jerarquización analítica mejorada para el desarrollar un sistema de criterios genéricos que incluye criterios duros y suaves en aspectos económicos, ambientales, sociales y tecnológicos para la evaluación de la sostenibilidad. Del mismo modo, en cuanto a la toma de decisiones, a partir de la generación de escenarios futuros en el sector energético, se investigan conceptos y enfoques para analizar, comparar y evaluar los diversos futuros energéticos desde un punto de vista epistemológico, para así llevar a cabo las comparaciones y evaluaciones de los dichos futuros energéticos, mediante el análisis de los componentes que se han utilizado en la construcción de los respectivos futuros, y en el proceso de su composición.

En cuanto a investigaciones relacionadas con la evaluación cualitativa de las estructuras encargadas de la toma de decisiones en las organizaciones, se enfoca en un tratamiento general que permite analizar la gama completa de arquitecturas organizacionales entre

formas extremas centralizadas y descentralizadas. Además, esta evaluación cualitativa permite a los diseñadores examinar la confiabilidad general de la estructura organizacional a medida que cambia el número de actores dentro de la organización.

Durante muchos años, los escenarios se han promovido como una técnica clave para la toma de decisiones e implementación de estrategias en entornos inciertos; sin embargo, la inclusión de todas las partes interesadas, y no sólo de los expertos, es algo que ha adquirido importancia para la toma de decisiones. Además de que el análisis de escenarios es un método probado para el apoyo en la toma de decisiones y la selección de estrategias en condiciones de incertidumbre, las investigaciones más recientes con respecto a la formulación de escenarios, tomando en cuenta la participación, recientemente se han enfocado en el desarrollo y análisis de escenarios transdisciplinarios, donde se genera un pequeño conjunto de los mismos, que pueden reflejar una complejidad sustancial y multidimensional, pero que son aptos para su uso dentro de la investigación participativa. Al ofrecer una plataforma para el diálogo transdisciplinario, los escenarios alteran los supuestos existentes, reintroducen cuestiones sociales y de justicia a los desafíos, lo que permite entender dichos escenarios transdisciplinarios como un método prospectivo de creación de sentido, lo que implica una forma de fomentar la discusión sobre el futuro y comprender las consecuencias de las actividades de hoy en las diversas realidades del mañana a partir de la perspectiva de las partes interesadas y/o afectadas, lo cual es crucial para construir una comprensión de eventos futuros.

Sin embargo, otras líneas de indagación que abordan la falta de estándares de calidad en la investigación transdisciplinaria, proponen guías como bases de calidad, que van dirigidas a tres grupos de actores: investigadores, administradores de programas y formuladores de políticas. Con esto se asegura la calidad del proceso de investigación transdisciplinaria.

Del mismo modo, los estudios en torno a este concepto determinan la transdisciplinariedad como un enfoque para la evaluación de iniciativas, es decir, el proceso transdisciplinario describe la naturaleza de las relaciones de colaboración entre investigadores y otras partes interesadas, y examina los factores que inhiben y facilitan dicha colaboración, además de que destacan los efectos que este enfoque podría tener en el futuro, en términos de construcción de redes, coproducción de conocimiento y desarrollo de prácticas innovadoras.

Otra de las líneas de investigación de la transdisciplinariedad incluye la distinción entre dos formas diferentes de dicha transdisciplinariedad: participativa y consultiva. Esta distinción se basa en la diferencia cualitativa entre investigaciones realizadas que incluyen a todo tipo de actores en igualdad de condiciones durante el proceso de producción de conocimiento (transdisciplinariedad participativa), o que actores externos a la academia respondan y reaccionen a la investigación realizada (transdisciplinariedad consultiva). Ambas formas cumplen los requisitos básicos de la transdisciplinariedad, pero difieren en cuanto a los desafíos involucrados, por lo que es necesario hacer una distinción entre ellas al discutir o evaluar la investigación.

Asimismo, se hace énfasis en aumentar la participación de actores externos al ambiente académico en los procesos de producción de conocimiento, llamado coproducción transdisciplinaria, si bien la coproducción transdisciplinaria aumenta la usabilidad de los resultados en términos de su relevancia y accesibilidad, que paradójicamente no asegura su anclaje en los respectivos contextos institucionales y políticos donde se produce el cambio social.

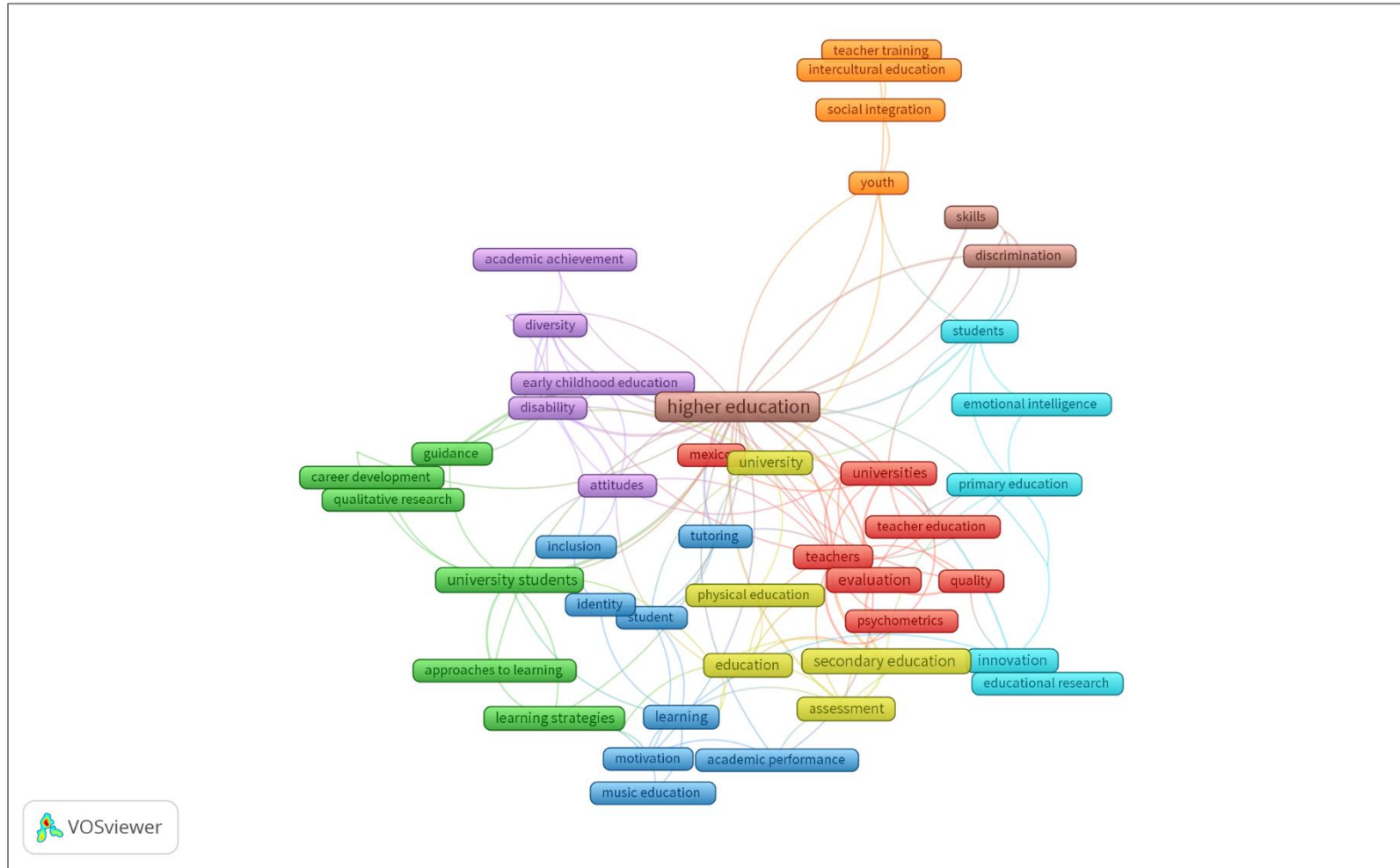
Educación

Otra de las áreas temáticas que se estudiaron en esta investigación es la de educación: se revisaron 316 documentos obtenidos de las revistas.

Al no contar con un número grande de elementos, no fue necesario filtrar las palabras clave. Sin embargo, no todos los clústeres generados fueron relevantes para alcanzar los objetivos de este trabajo. Por ello, las visualizaciones se enfocaron en aquellos clústeres cuyo concepto de evaluación estuviera incluido.

Los clústeres generados a partir de los 316 documentos se muestran en el *Mapa 12*, donde cada uno se representa con un color. Asimismo, para identificar la ubicación de los diferentes conceptos del flujo de información en el área de educación a lo largo del periodo de tiempo en estudio, se decidió elaborar la visualización superpuesta mostrada en el *Mapa 13*. De igual forma que en el caso del área de sistemas y planeación, estos mapas por sí solos no arrojan información suficiente para establecer alguna aseveración. Sin embargo, al hacer el mismo ejercicio de centralizar el ítem *evaluation*, se muestran con mayor detalle los conceptos relacionados con éste.

Mapa 12. Mapa de red de coocurrencia de palabras claves de las revistas de educación.



Fuente: Elaborado a partir de información de las diferentes bases de datos, utilizando VOSviewer (fecha de consulta: 12/02/2020).

A partir de esta relación entre conceptos, se encontró que las investigaciones referentes a evaluación en el área de educación, en su mayoría, se refieren a la evaluación del desempeño tanto de estudiantes como de profesores de los niveles básicos, medio superior y superior, gracias a un enfoque cualitativo y otro cuantitativo, mediante la aplicación de herramientas como entrevistas, cuestionarios, grupos de control, entre otros, con el objetivo de fortalecer la calidad educativa. Del mismo modo, en cuanto a la evaluación de programas educativos, las investigaciones relacionadas a éstos presentan un enfoque experimental, es decir, se caracterizan por el diseño de la evaluación, procedimientos de recolección de datos, población y muestra, así como de análisis de datos, e instrumentos de recolección de datos.

También se encontró que los trabajos relacionados con este concepto abordan el tema de la evaluación participativa, donde no sólo se toma en cuenta la opinión del estudiante, sino también la participación de los profesores, gracias a técnicas participativas como la autoevaluación, la evaluación por pares y la coevaluación. Sin embargo, los resultados de estas investigaciones arrojan que el uso de estas técnicas participativas de evaluación es poco común en las universidades, además de que no existen procesos de formación tanto para los profesores como los alumnos, lo que permitiría mejoras en el aprendizaje independiente y estratégico de los estudiantes.

CONCLUSIONES

A través de los diferentes capítulos de este trabajo de investigación se propuso el objetivo principal de establecer un estado del arte sobre los nuevos estudios técnicos, metodológicos y conceptuales, y líneas de investigación que de Evaluación se han publicado desde el 2010 al 2019.

A partir de la visualización y discusión anteriores, se pueden establecer las siguientes conclusiones, las cuales presentan reflexiones y, los vacíos y necesidades para investigar sobre el campo de Evaluación.

Conclusión 1. La estructura intelectual en la que se sustentan las investigaciones publicadas en los journals de Evaluación, principalmente está influenciada por Michael Patton, así como a J. Bradley Cousin, Jennifer C. Greene, Donald T. Campell, Thomas D. Cook, Carol Hirschon Weiss, William R. Shadish, Hallie Preskill, Abraham Wandersman, Lee J. Cronbach, Michael Scriven y Melvin M. Mark. Es decir, las publicaciones del 2010 al 2019 citaron con mayor frecuencia a los autores mencionados,

los cuales representan de una manera más relevante la base del conocimiento de las publicaciones de Evaluación.

Conclusión 2. Durante la revisión, se encontró que los enfoques predominantes de evaluación son el cualitativo, cuantitativo y experimental a través de la aplicación de técnicas como entrevistas, cuestionarios, costo-beneficios, grupos de control, entre otros. Es decir, en un periodo de 10 años, estos enfoques han servido para llevar a cabo evaluaciones, en específico de programas sociales, educativos, en la toma de decisiones, desempeño de estudiantes y maestros, en la selección de alternativas, etc.

Conclusión 3. En cuanto al área de Sistemas y Planeación, se encontró que los documentos relacionados con Evaluación, están bajo en contexto de tecnología e innovación. Y por lo general, recurren a métodos y técnicas cualitativas y cuantitativas que incluyen aspectos duros y suaves, como es el caso de las entrevistas y la jerarquización analítica, pese a las limitaciones de esta última.

Conclusión 4. A partir de la construcción de los mapas bibliométricos, se pudieron identificar aquellas líneas de investigación o vacíos y necesidades que a lo largo de 10 años no se han desarrollado sobre Evaluación, y que hasta el 2019 no se cuenta con información suficiente o documentada del seguimiento de estas líneas, las cuales son:

- Diseño de evaluaciones con base en la práctica. Es decir, es necesario que los evaluadores documenten toda la evidencia de las actividades de campo para que, con base en esto, se diseñen evaluaciones apegadas a la vida real, y así lograr la intersección de la teoría y la práctica.

- Lineamientos y bases para la profesionalización de la Evaluación, así mismo, para considerar a la Evaluación como una transdisciplina.
- Diseño de evaluaciones multisitio que incluyan las características y propiedades locales en donde se lleva a cabo la ejecución del programa o proyecto.
- Diseño y aplicación de metodologías y técnicas para identificar nuevas partes interesadas, y no solo la parte experta, y así llevar a cabo la construcción de escenarios de evaluación y toma de decisiones. Y con esto, realizar una evaluación completamente participativa.
- Evaluación participativa mediante la construcción de escenarios transdisciplinarios donde se involucren actores externos a una sola disciplina, es decir, comprender las consecuencias de las decisiones tomadas vistas desde diferentes enfoques.

BIBLIOGRAFÍA

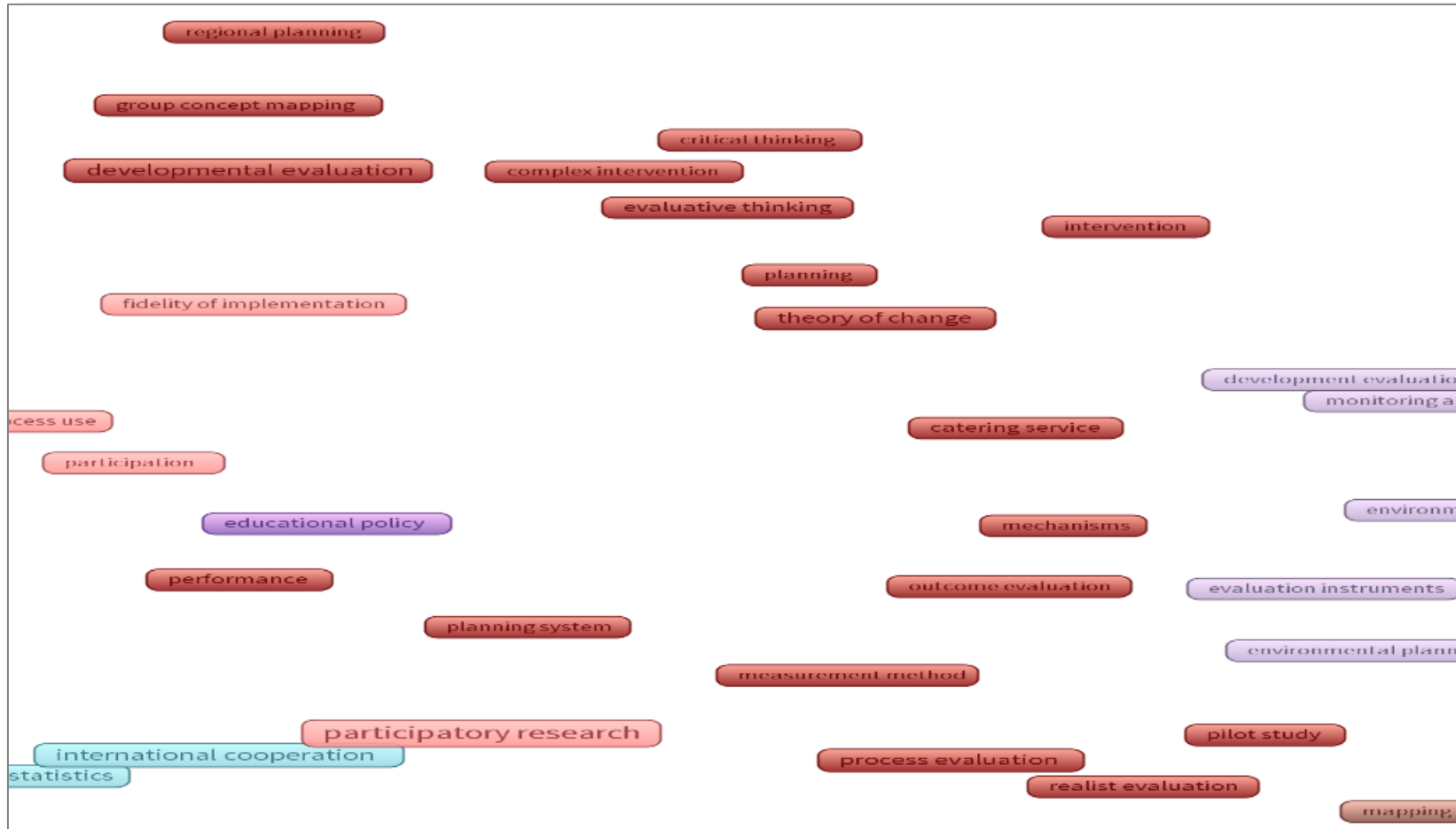
- Alcaraz, N. (2015). Aproximación Histórica a la Evaluación Educativa: De la Generación de la Medición a la Generación Ecléctica. *Revista Iberoamericana de Evaluación Educativa*, 11-25.
- Alkin , M., Vo, A., & Christie, C. (2012). The evaluator's role in valuing: Who and with whom. *New Directions for Evaluation*, 29-41.
- Alkin, M., & King , J. (2016). The Historical Development of Evaluation Use. *American Journal of Evaluation*, 568-579.
- Alkin, M., Vo, A., & Christie, C. (2012). The evaluator's role in valuing: Who and with whom. *New Directions for Evaluation*, 29-41.
- Araujo, S. M. (2014). La evaluación y la Universidad en Argentina: políticas, enfoques y prácticas. *Revista de la Educación Superior*, 57-77.
- Archibald , T., Sharrock , G., Buckley , J., & Cook, N. (2016). Assumptions, conjectures, and other miracles: The application of evaluative thinking to theory of change models in community development. *Evaluation and Program Planning*, 119-127.
- Bryson, J., Patton, M., & Bowman, R. (2011). Working with evaluation stakeholders: A rationale, step-wise approach and toolkit. *Evaluation and Program Planning*, 1-12.
- Buckley , J., Archibald , T., Hargraves, M., & Trochim W. (2015). Defining and Teaching Evaluative Thinking: Insights From Research on Critical Thinking. *American Journal of Evaluation*, 375-388.
- Cabrera, D., Colosi, L., & Lobdell, C. (2008). Systems thinking. *Evaluation and Program Planning*, 299-310.
- Chapman , H., Murphy, N., & Patton, M. (2017). A Guiding Typology for Site Visits. *New Directions for Evaluation*, 11-19.

- Christie , C., Lemire , S., & Inkelas, M. (2017). Understanding the Similarities and Distinctions Between Improvement Science and Evaluation. *New Directions for Evaluation*, 11-21.
- Christie, C., & Lemire, S. (2019). Why Evaluation Theory Should Be Used to Inform Evaluation Policy. *Canadian Journal of Program Evaluation*, 395–413.
- Christie, C., Quiñones , P., & Fierro, L. (2014). Informing the Discussion on Evaluator Training: A Look at Evaluators' Course Taking and Professional Practice. *American Journal of Evaluation*, 274-290.
- Conner, R. (2012). Reflections over 25 years: Evaluation Then, Now, and Into the Future. *The Canadian Journal of Program Evaluation*, 125-136.
- Coryn, C. L., Wilson, L. N., Westine, C. D., Hobson, K. A., Oseki, S., Fiekowsky, E. L., . . . Schröter, D. C. (2017). A Decade of Research on Evaluation: A Systematic Review of Research on Evaluation Published Between 2005 y 2014. *American Journal of Evaluation*, 329-347.
- Coryn, C., Oseki, S., Wilson, L. N., Greenman II, G. D., Schröter, D. C., Hobson, K. A., . . . Vo, A. T. (2016). Does Research on Evaluation Matter? Findings From a Survey of American Evaluation Association Members and Prominent Evaluation Theorists and Scholars. *American Journal of Evaluation*, 159-173.
- Donaldson , S. (2017). Empowerment evaluation: An approach that has literally altered the landscape of evaluation. *Evaluation and Program Planning*, 136-137.
- Donaldson, S. (2019). Where do we stand? Recent AEA member views on professionalization. *Evaluation and Program Planning*, 152-161.
- Fierro , L., & Christie , C. (2017). Evaluator and Program Manager Perceptions of Evaluation Capacity and Evaluation Practice. *American Journal of Evaluation*, 376-392.
- Gálvez, C. (2016). Visualización de las principales líneas de investigación en salud pública: un análisis basado en mapas bibliométricos aplicados a la Revista Española de Salud Pública (2006-2015). *Revista Española de Salud Pública*, e1-e10.
- Guba, E., & Yvonna, L. (1989). *Fourth Generation Evaluation*. California: Sage Publications.
- Gullickson , A., King , J., LaVelle, J., & Clinton , J. (2019). The current state of evaluator education: A situation analysis and call to action. *Evaluation and Program Planning*, 20-30.
- Guzmán Sánchez, M. V., & Trujillo Cancino, J. (2013). Los mapas bibliométricos o mapas de la ciencia: una herramienta útil para desarrollar estudios métricos de información. *Revistas UNAM*, 95-108.
- House, E. (2014). Origins of the ideas in Evaluating with Validity. *New Directions for Evaluation*, 9-15.
- House, E. (2019). Evaluation With a Focus on Justice. *New Directions for Evaluation*, 61-72.
- Julnes , G., & Bustelo, M. (2017). Professional Evaluation in the Public Interest(s). *American Journal of Evaluation*, 540-545.

- Julnes, G. (2012). Managing valuation. *New Directions for Evaluation*, 3-15.
- Lawrenz, F., King, J., & Ooms, A. (2011). The role of involvement and use in multisite evaluations. *New Directions for Evaluation*, 232-237.
- Leviton, L. C., & Gutman, M. A. (2010). Overview and Rationale for the Systematic Screening and Assessment Method. *New Directions for Evaluation*, 7-31.
- Muñoz, G. A. (2007). Un nuevo paradigma: "la quinta generación de la evaluación". *Lauros*, 158-198.
- Murillo, F. J., & Román, M. (2010). Retos en la Evaluación de la Calidad de la Educación en América Latina. *Revista Iberoamericana de Educación*, 97-120.
- Patton, M. (2015). Evaluation in the Field: The Need for Site Visit Standards. *American Journal of Evaluation*, 444-460.
- Patton, M. (2017). Revised Site-Visit Standards: A Quality-Assurance Framework. *New Directions for Evaluation*, 83-102.
- Patton, M. (2018). Evaluation Science. *American Journal of Evaluation*, 183-200.
- Patton, M. Q. (2012). Incomplete Successes. *The Canadian Journal of Program Evaluation*, 151-163.
- Sánchez, G. (1994). *Una propuesta Teórico-Metodológica para la Evaluación de Programas Sociales: Enfoque a programas académicos universitarios*. México.
- Shadish, W., Cook, T., & Leviton, L. (1991). *Foundations of program evaluation: theories of practice*. SAGE Publications, Inc.
- Stufflebeam, D., Madaus, G., & Kellaghan, T. (2002). *Evaluation Models. Viewpoints on Educational and Human Services Evaluation* (Segunda ed.). KLUWER ACADEMIC PUBLISHERS.
- Suchman, E. (1967). *Evaluative Research*. Russell Sage Foundation.
- Wasserman, D. (2010). Using a systems orientation and functional theory to enhance theory-driven human service program evaluations. *Evaluation and Program Planning*, 67-80.

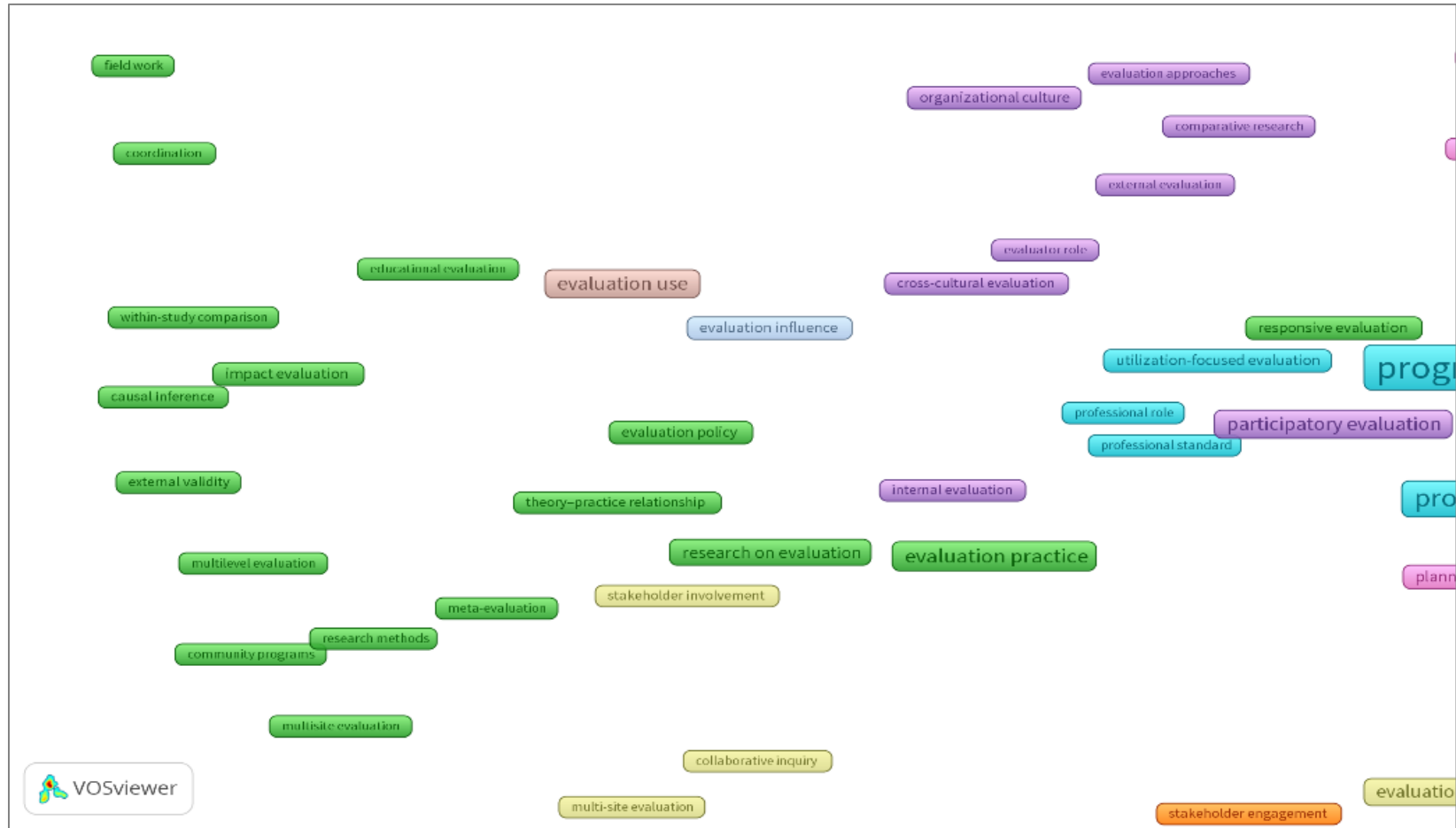
ANEXO A

Mapa 15. Clúster 1 (rojo). Developmental Evaluation.



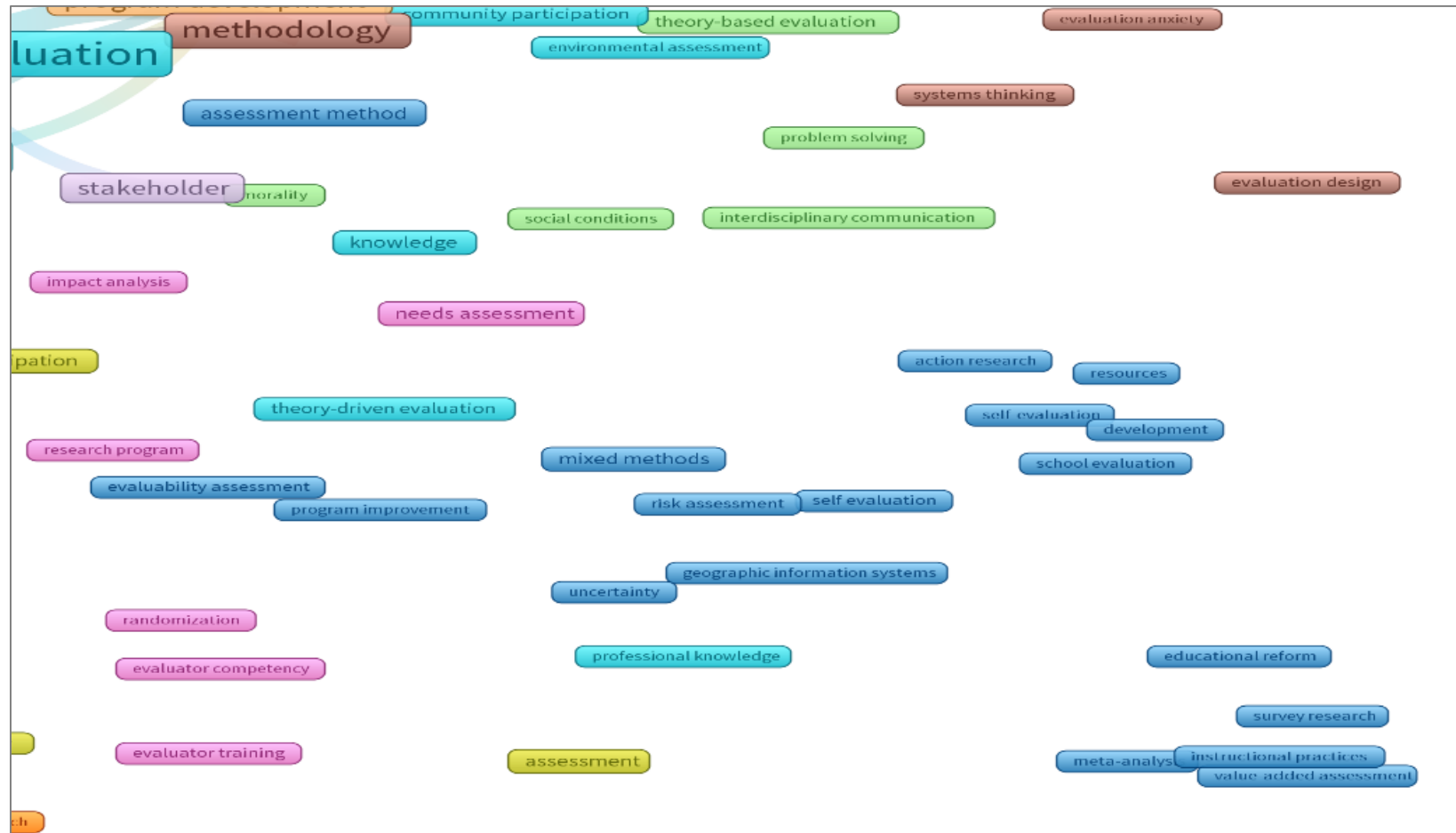
Fuente: Elaborado a partir de información de las diferentes bases de datos, utilizando VOSviewer (fecha de consulta: 12/02/2020).

Mapa 16. Clúster 2 (verde). Evaluation Practice.



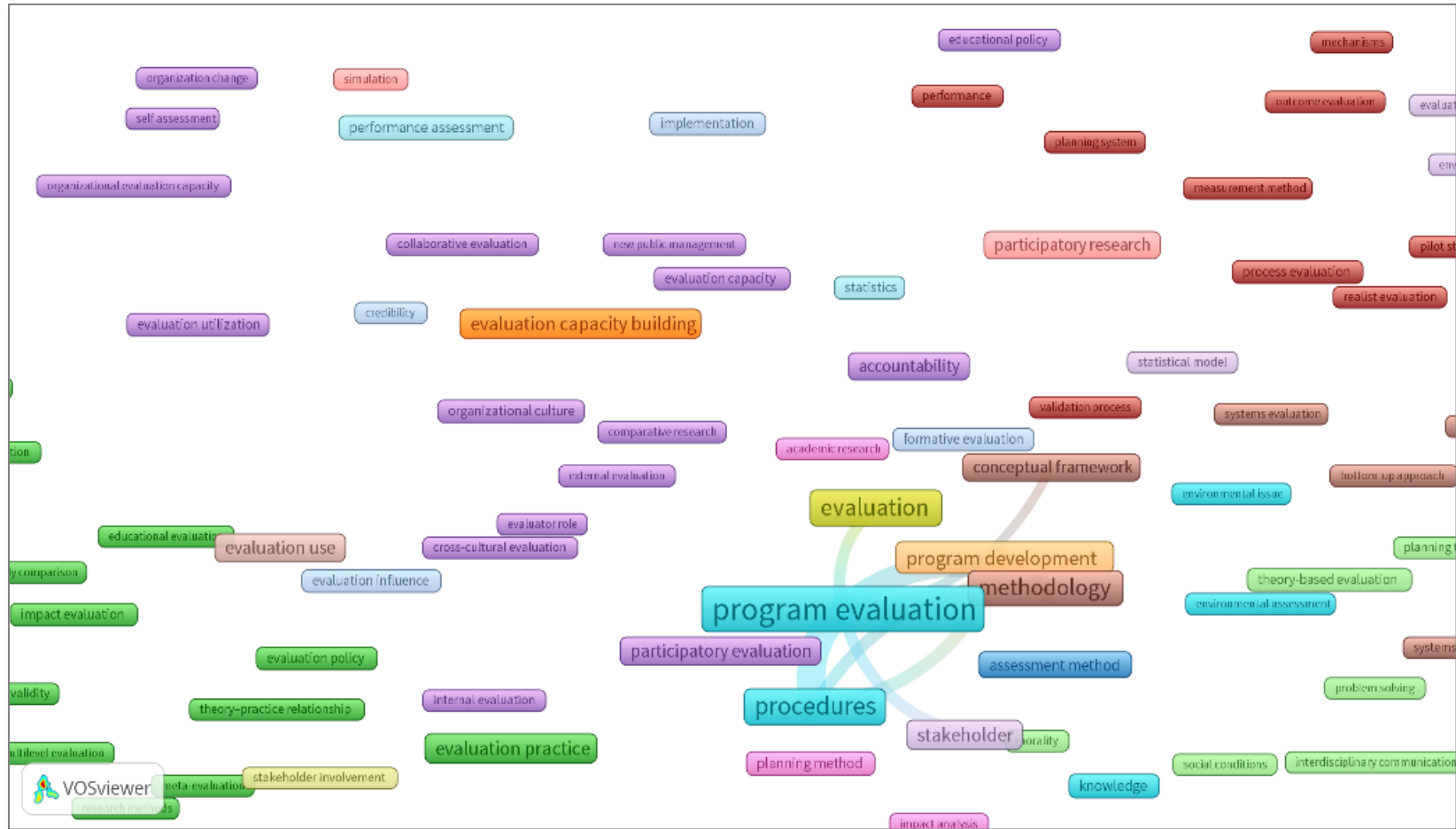
Fuente: Elaborado a partir de información de las diferentes bases de datos, utilizando VOSviewer (fecha de consulta: 12/02/2020).

Mapa 17. Clúster 3 (azul fuerte). Assessment Method.



Fuente: Elaborado a partir de información de las diferentes bases de datos, utilizando VOSviewer (fecha de consulta: 12/02/2020).

Mapa 19. Clúster 5 (morado). Participatory Evaluation.



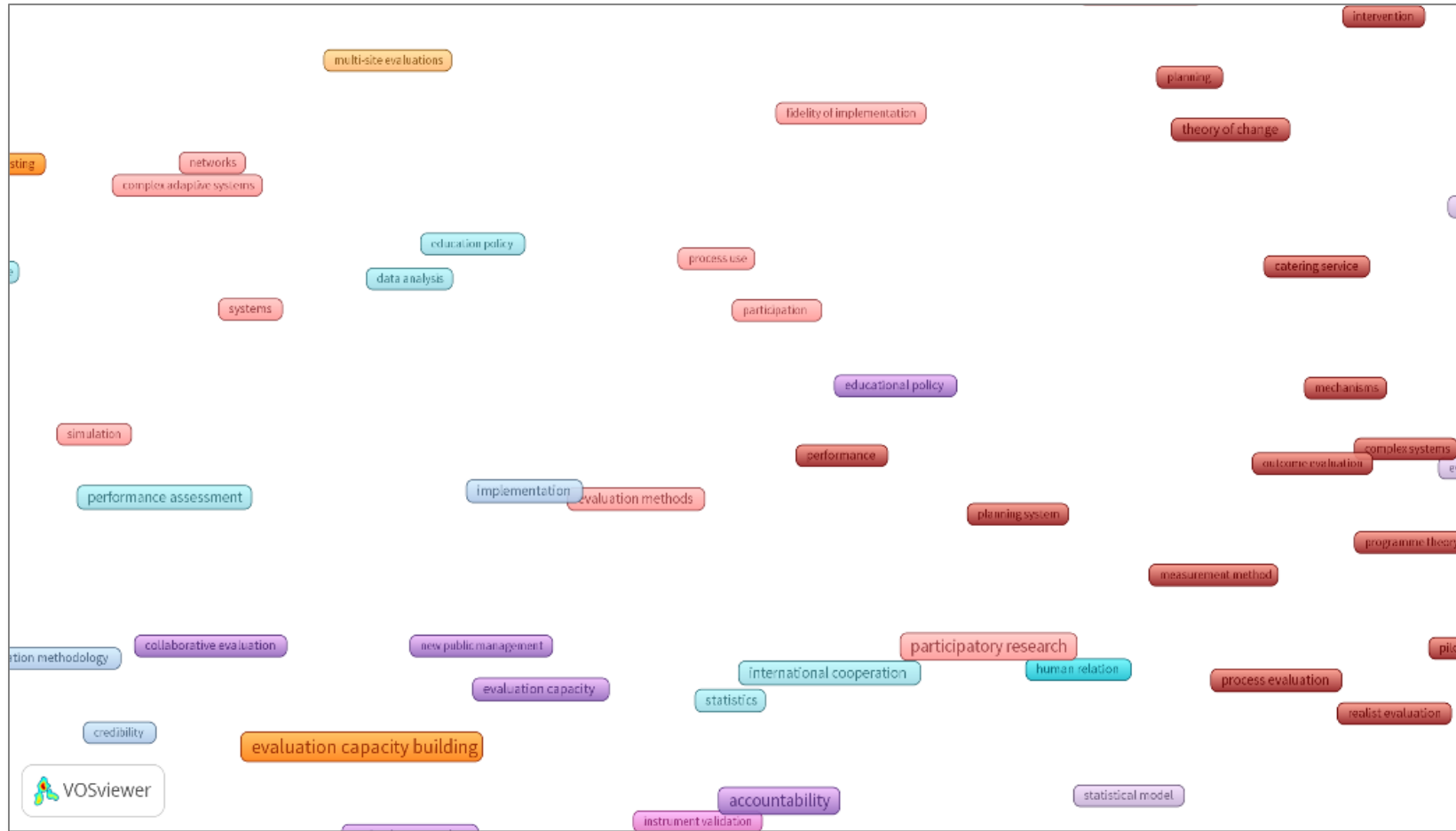
Fuente: Elaborado a partir de información de las diferentes bases de datos, utilizando VOSviewer (fecha de consulta: 12/02/2020).

Mapa 22. Clúster 8 (marrón). Methodology.



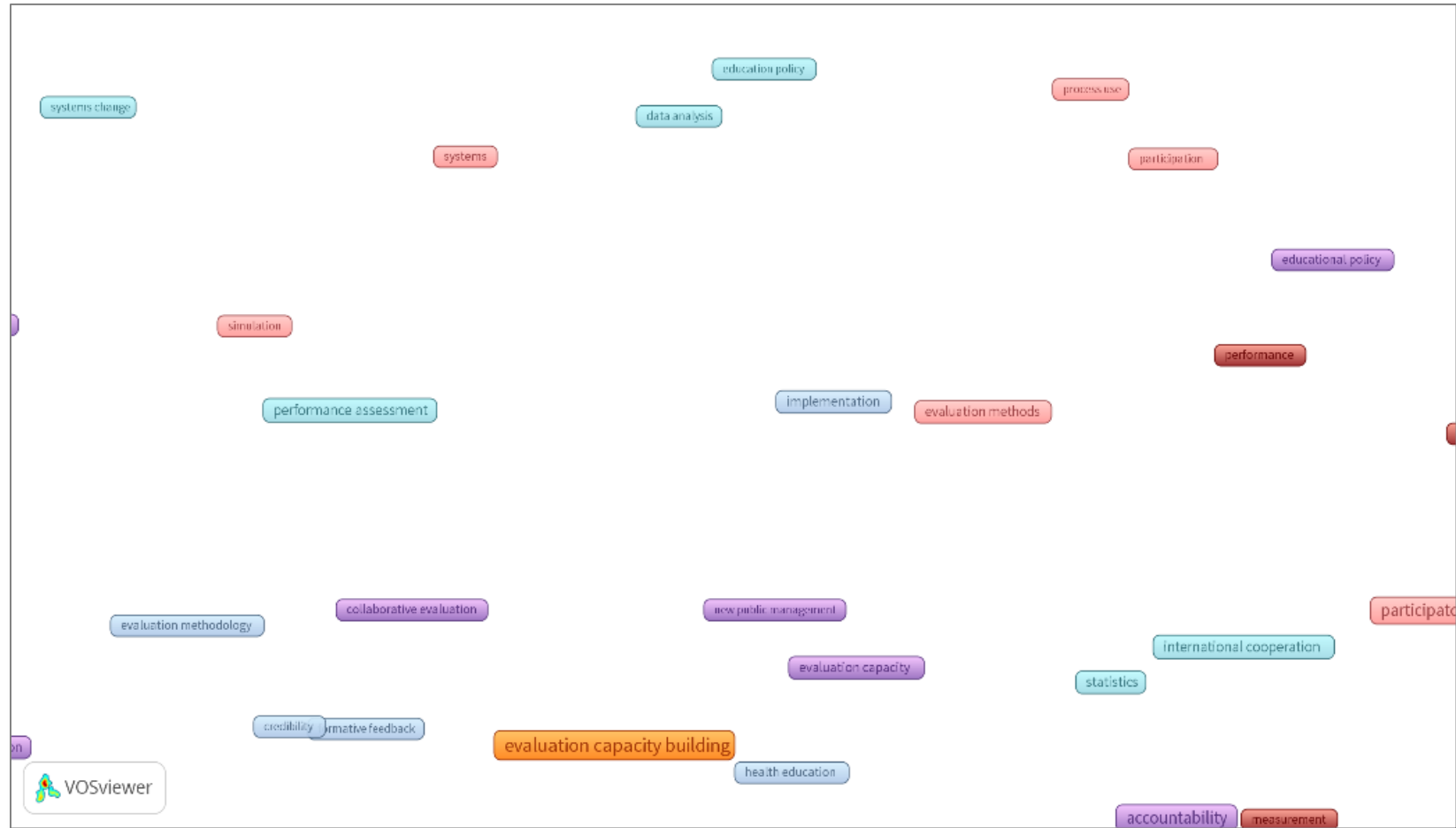
Fuente: Elaborado a partir de información de las diferentes bases de datos, utilizando VOSviewer (fecha de consulta: 12/02/2020).

Mapa 24. Clúster 10 (salmón). Participatory Research.



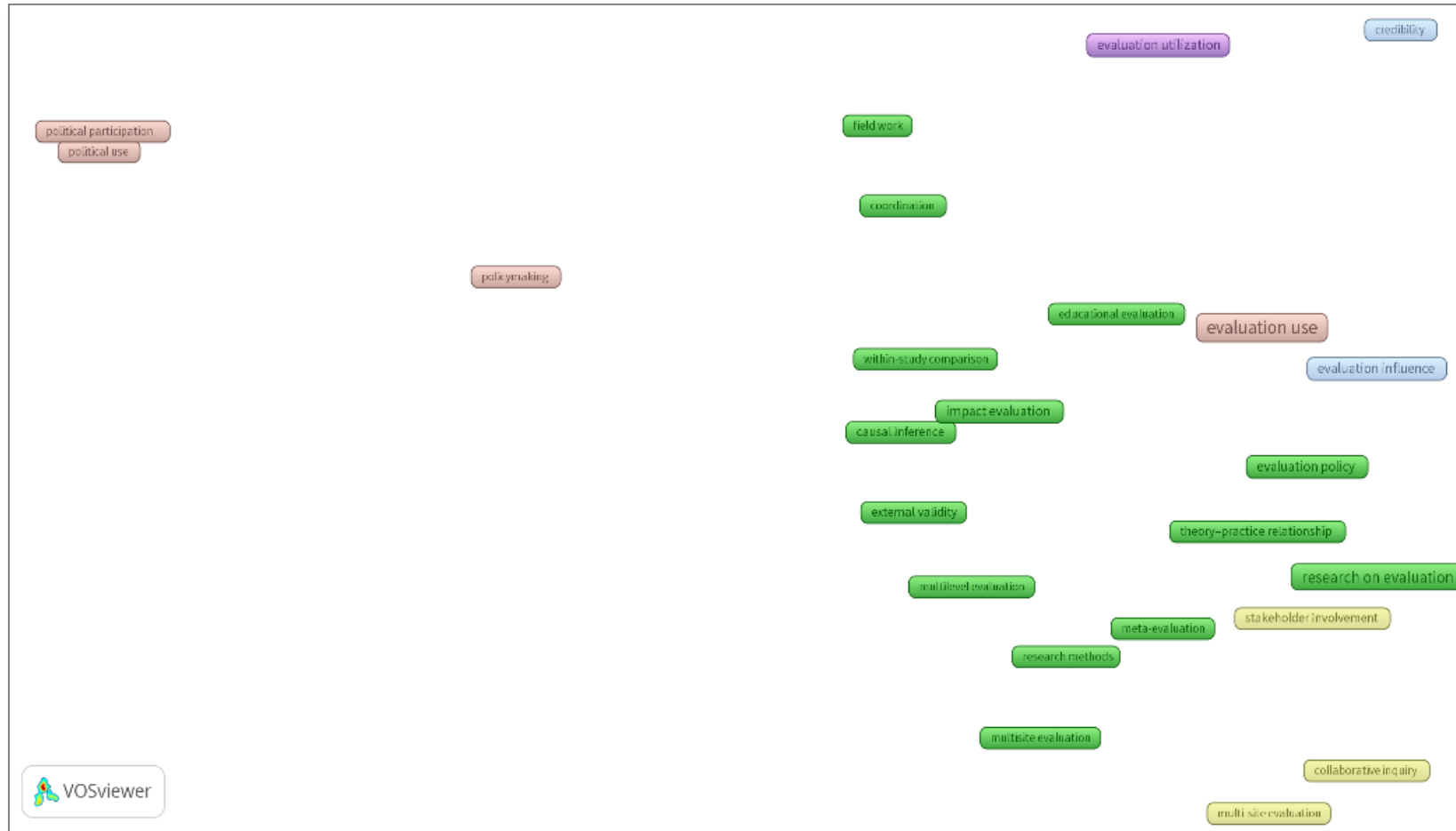
Fuente: Elaborado a partir de información de las diferentes bases de datos, utilizando VOSviewer (fecha de consulta: 12/02/2020).

Mapa 29. Clúster 15 (azul cielo). Performance Assessment.



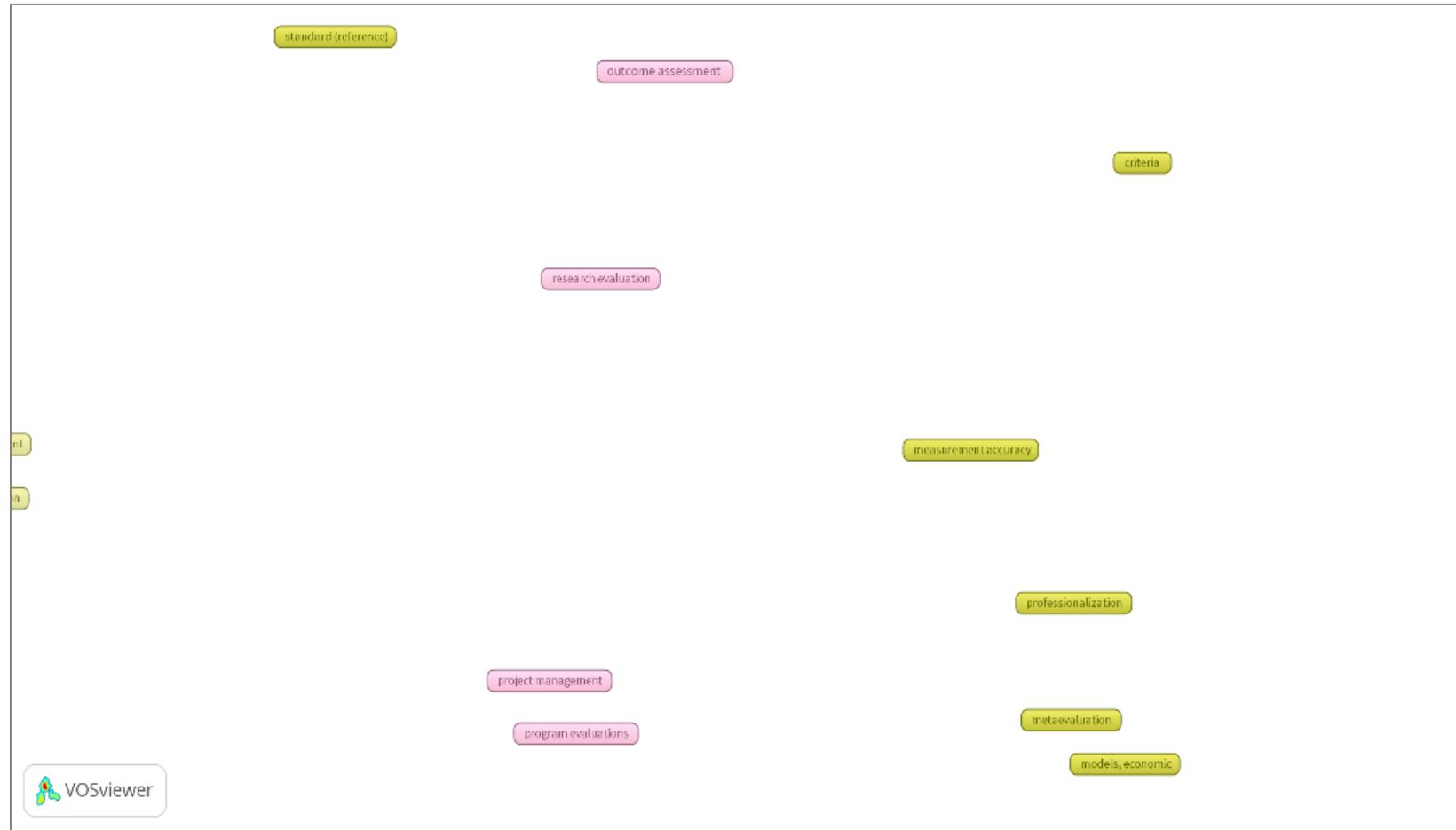
Fuente: Elaborado a partir de información de las diferentes bases de datos, utilizando VOSviewer (fecha de consulta: 12/02/2020).

Mapa 31. Clúster 17 (marrón). Evaluation Use.



Fuente: Elaborado a partir de información de las diferentes bases de datos, utilizando VOSviewer (fecha de consulta: 22/10/2019).

Mapa 32. Clúster 18 (rosa). Outcome Assessment.



Fuente: Elaborado a partir de información de las diferentes bases de datos, utilizando VOSviewer (fecha de consulta: 12/02/2020).

ANEXO B

A continuación, se presenta una breve guía para quienes empiezan a usar el programa VOSviewer. El programa puede descargarse de <http://www.vosviewer.com> y ser usado libremente para cualquier propósito.

VOSviewer es una herramienta de software para construir y visualizar redes bibliométricas. Estas redes pueden incluir, por ejemplo, revistas, investigadores o publicaciones individuales, y pueden construirse sobre la base de la citación, el acoplamiento bibliográfico, la cocitación o las relaciones de coautoría. VOSviewer también ofrece una funcionalidad de minería de textos que puede utilizarse para construir y visualizar redes de coocurrencia de términos importantes extraídos de un cuerpo de literatura científica.

La funcionalidad del VOSviewer puede ser resumidas de la siguiente manera¹:

- Creación de mapas basados en datos de la red. Se puede crear un mapa basado en una red que ya está disponible, pero también es posible construir uno de la red. VOSviewer puede utilizarse para construir redes de publicaciones científicas, revistas científicas, investigadores, organizaciones de investigación, países, palabras clave, o condiciones. Los elementos de estas redes pueden estar conectados por enlaces de coautoría, coocurrencia, citación, acoplamiento bibliográfico o cocitación. Para construirlos, se pueden descargar bases de datos de los repositorios como Web of Science, Scopus, Dimensions, PubMed, RIS, o Crossref JSON se pueden utilizar los archivos.
- Visualizar y explorar mapas. VOSviewer proporciona tres visualizaciones de un mapa: La visualización de la red, la visualización de la superposición, y la densidad visualización. La función de zoom y desplazamiento que permite explorar un mapa con detalles completos, lo cual es esencial cuando se trabaja con mapas grandes que contienen miles de artículos.

¹ <https://universoabierto.org/2020/02/18/vosviewer-es-una-herramienta-de-software-para-construir-y-visualizar-redes-bibliometricas/>

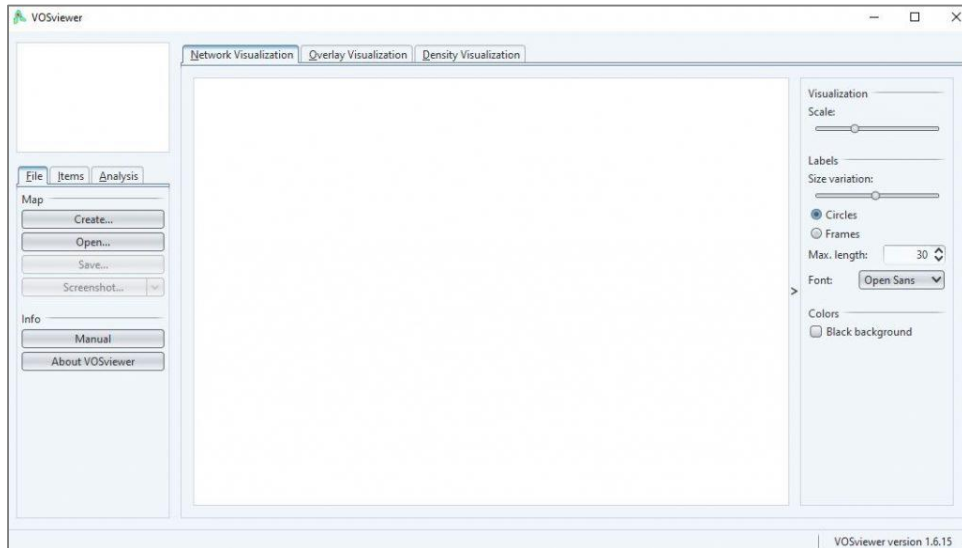
Estado del Arte del Campo Disciplinario de la Evaluación

Es importante mencionar que antes de empezar a usar el programa, ya debemos tener descargada nuestra base de datos. En este caso, se descargó de Scopus.

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100																												
1	Authors	Author(s) Title	Year	Source title	Volume	Issue	Art. No.	Page start	Page end	Page count	Cited by	DOI	Link	Affiliation	Abstract	Author Key	Index Keyw	Molecular	Chemical	Tradname	Manufacturer	Funding De	Funding Tc	Funding Tr	Reference	Correspond	Editors	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100

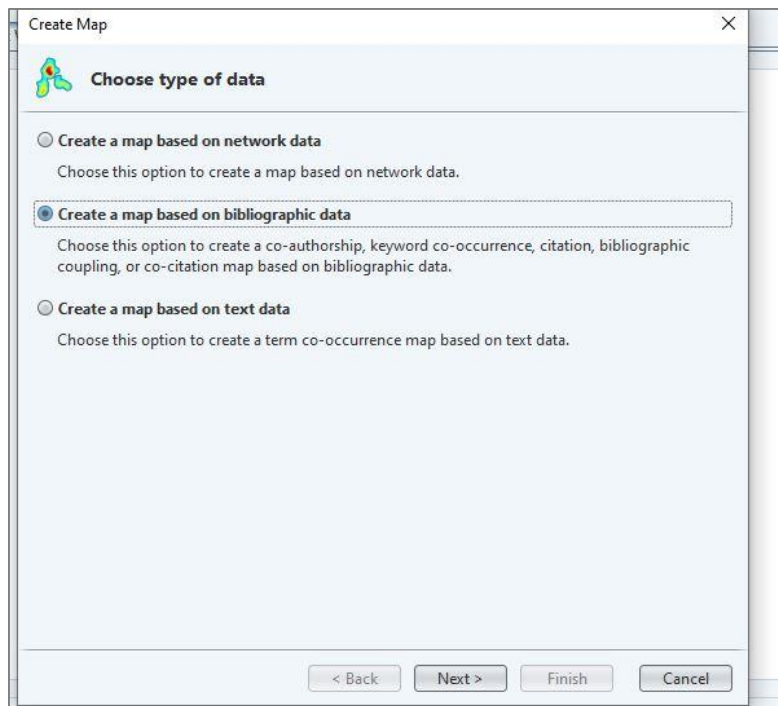
El siguiente procedimiento es únicamente ilustrativo, ya que no corresponde a un estudio en cuestión:

- 1 Inicie VOSviewer. Una vez descargado, descomprima la carpeta y dé doble clic en el ícono de VOSviewer). Debido a que el programa ha sido desarrollado por Java, no es necesario instalarlo.

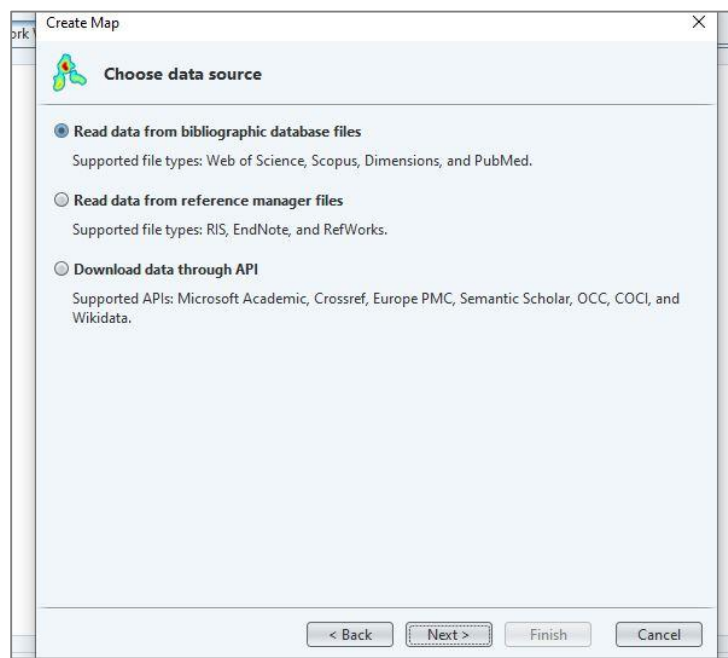


² <http://www.revistaotlet.com/tips-cesar-limaymanta-mapeo-cientifico-con-vosviewer/>

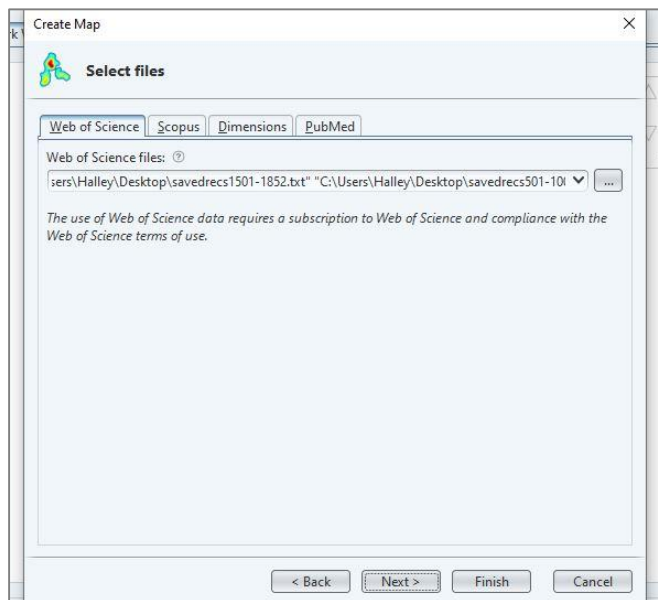
2 Presione el botón *Create* (lado izquierdo) y seleccione el botón de opción *Create a map based on bibliographic data* y presione el botón *Siguiente*.



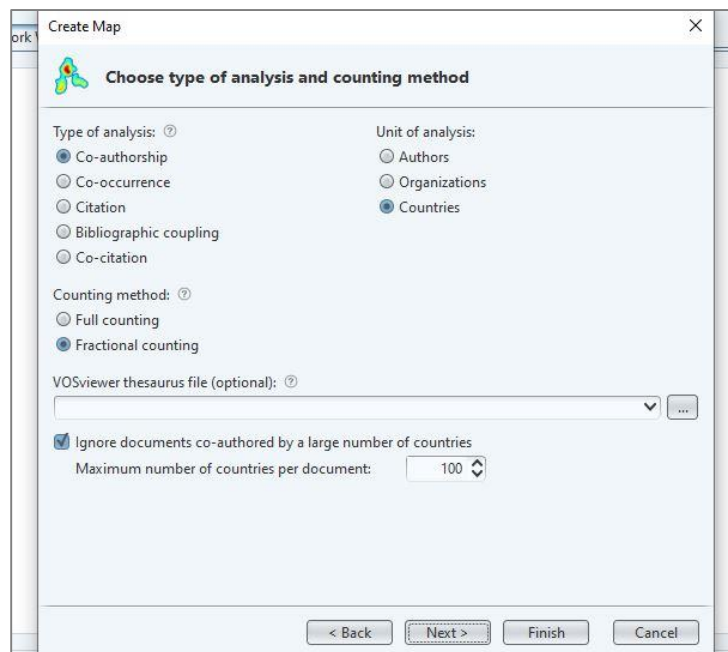
3 Seleccione la opción *Read data from bibliographic database files* y presione *Siguiente*.



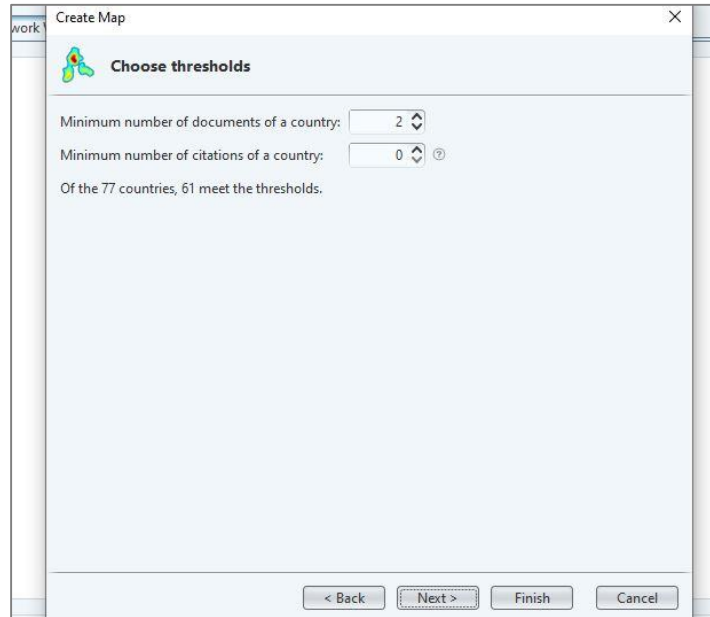
4 Vaya a la pestaña Scopus o cualquiera de donde haya descargado la base de datos para abrir el cuadro de diálogo y seleccione los metadatos descargados desde Scopus. La tecla Shift se puede utilizar para seleccionar varios archivos. Se selecciona los xxx registros obtenidos. Presione el botón OK y luego presione Siguiente.



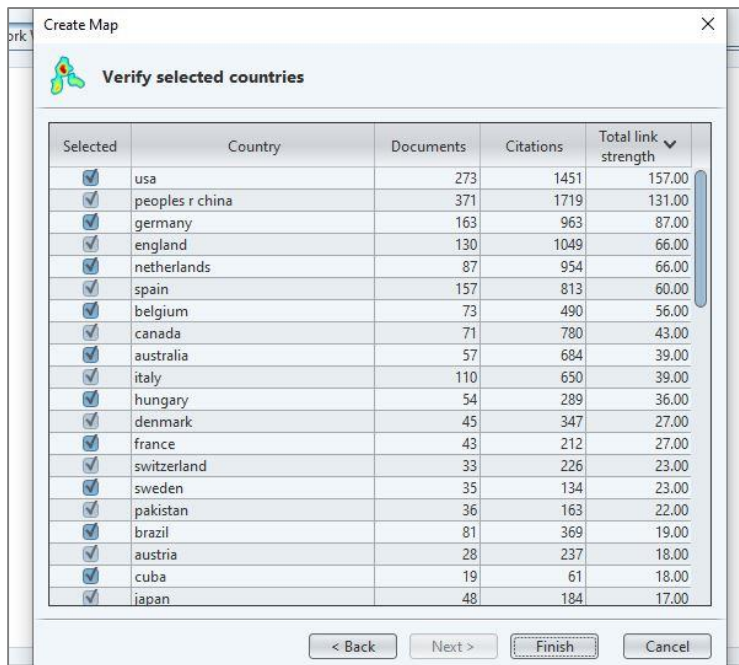
5 Seleccione la opción del tipo y unidad de análisis, y el método de conteo que necesite, y presione el botón Siguiente.



6 VOSviewer solicita el número mínimo de publicaciones que desee tener en su análisis para la construcción del mapa de red. Luego, presione el botón *Siguiente*. En la nueva ventana, presione también *Siguiente*.



7 VOSviewer enumera los resultados incluidos en la red y ofrece la posibilidad de eliminar aquellos resultados no deseados. Por último, presione *Finish*. En este paso, se generan los clústeres, de los cuales algunos llegan a ser bastante pequeños. Puede resultar útil reducir el número de clústeres de esta manera:



8 Para este ejemplo, se muestra en la siguiente figura la visualización de coautoría entre países. Cada círculo representa un país y su tamaño refleja la cantidad de documentos publicados en la revista *Scientometrics*. La cercanía o lejanía de un país a otro refleja la fuerza de coautoría que ejerce cada país. Es decir, cuanto más cerca se encuentra un país de otro, mayor será la fuerza de coautoría entre ellos. Mientras que países que se ubican lejos uno del otro, menor será la relación de coautoría. Los colores indican clúster de países que están relativamente relacionados entre sí.

