

Capítulo 2

Análisis de la percepción sobre el SC&CTI del departamento del Valle del Cauca

Doi:

10.25100/peu.689.cap2

Autores:

Jhon James Mora Rodríguez.
Universidad Icesi

 0000-0001-5499-5400

Diana Yaneth Herrera Duque
Universidad Icesi

 0000-0001-5110-0683

Lya Paola Sierra
P. Universidad Javeriana Cali

 0000-0002-8909-8977

Kiara Fernanda Ramírez Orozco
Universidad Icesi

 0000-00003-0319-9652

Pavel Vidal Alejandro
P. Universidad Javeriana Cali

 0000-0001-8278-3122

Juan Tomas Sayago
Universidad Icesi

 0000-0002-7238-124X

Julián Durán Peralta
Universidad Santiago de Cali

 0000-0002-8252-729X

Elizabeth Aponte Jaramillo
U. Autónoma de Occidente

 0000-0003-0820-5679

Paola Andrea Garizado Román
U. Autónoma de Occidente

 0000-0001-6420-262X

Lilian Andrea Carrillo Rodríguez
U. Autónoma de Occidente

 0000-0001-9045-2905

Henry Caicedo Asprilla
Universidad del Valle

 0000-0003-1839-7061

Las metodologías de análisis basados en minería de texto son el proceso de extraer intenciones o intuiciones de los textos (Kwartler, 2017). En la práctica, la minería de textos se usa para identificar acciones a través de mensajes en redes sociales. Así, la información en textos es relevante por los múltiples usos que puede tener y la utilidad que tiene dentro de las redes sociales, la información de organizaciones, la digitalización de archivos, las nuevas tecnologías que permiten la transcripción automática y capturar diversas opiniones, y las fuentes textuales, las cuales continúan creciendo en cantidad y complejidad (Kwartler, 2017).

En el marco de estas aplicaciones, el objetivo de este capítulo consiste en desarrollar un análisis textual y una aproximación de sentimientos modificada a las entrevistas² realizadas a expertos en temas de C&CTI en el marco del Proyecto "Fortalecimiento del Sistema de C&CTI del Valle del Cauca: hacia una Economía del conocimiento." Este análisis permite establecer las tendencias en las opiniones de manera cuantitativa.

Estudios recientes como los trabajos de Wehnert *et al.* (2018) y Bellstam *et al.* (2020) han analizado la innovación de manera textual, el enfoque de estos trabajos ha sido sobre los reportes de actividades y otros registros públicos de las empresas. Se observan diferentes tendencias de investigación que se correlacionan con mediciones de éxito financiero y empresarial. Estos modelos logran argumentar una relación entre los procesos de innovación y los resultados económicos de las firmas.

Este capítulo se estructura de la siguiente forma: en primer lugar, se define la metodología a utilizar en el análisis textual; en segundo lugar se definen las preguntas utilizadas de la entrevista semiestructurada para la realización del análisis; en tercer lugar se describen los resultados alcanzados; y finalmente se presentan las conclusiones.

² Ver capítulo 1.

Metodología

Las técnicas utilizadas en este capítulo están relacionadas al análisis de minería de textos, las cuales realizan una caracterización y comprensión de los mensajes. Estos procesos permiten extraer o aproximarse a la idea central expresada por las personas en los textos. Establecer las ideas y los conceptos expresados por las personas permite encontrar tendencias y semejanzas, al mismo tiempo permite realizar un diagnóstico sobre las ideas centrales más comunes y repetidas entre las personas.

Existen dos tipos de procesos en minería de textos: El primero llamado “bolsa de palabras” trata cada palabra o grupos de palabras como características del documento y el segundo “emparejamiento sintáctico”, el cual es basado en la sintaxis de palabras y a partir de esta construir los componentes de la oración para definir el sentido. Cada uno tiene ventajas y beneficios, “bolsa de palabras” es relativamente sencillo y operacionalmente factible. Este método puede ser readecuado utilizando diferentes características. El método de “emparejamiento sintáctico” tiene como ventaja que el procesamiento y análisis es más rico en atributos, pero su procesamiento es más difícil de automatizar.

El método de análisis se divide en 5 etapas: 1) descarga y organización de la información, 2) organización, 3) extracción de características, 4) análisis y 5) organización del resultado analítico, ver Figura 2.1.

La primera etapa consiste en obtener la información a utilizar. La segunda consiste en organizar la información. En esta etapa se separan los textos como tokens o palabras para analizar, también se debe filtrar la información y retirar los conectores, números y otros elementos como links o hashtags. La tercera parte del análisis es la extracción de características, en esta etapa se toman las palabras restantes y se organizan para establecer las posibles conexiones entre palabras o los n-gramas. Esta etapa también se compara con el diccionario de palabras para clasificar el sentido de la frase. Esta clasificación

se lleva a cabo utilizando la regla de Bayes³ (ver ecuación 1). La cuarta etapa el análisis incluye analizar, clasificar y establecer relaciones en los discursos estudiados. Esta etapa utiliza técnicas de frecuencia y promedios de las características extraídas. La quinta etapa denominada como la organización del resultado analítico, incluye herramientas gráficas como diagramas de barras y nube de palabras. La nube de palabras – *Wordcloud* es una representación gráfica que se estructura a partir de las frecuencias de las palabras que surgen en un discurso, entre mayor sea la repetición de una palabra, mayor será el tamaño de su visualización, de modo que, entre menor sea su frecuencia, más pequeña será la visualización. Las nubes de palabras permiten captar visualmente las palabras, temáticas y expresiones más representativas en los discursos, identificando con ello las tendencias y discusiones centrales en la información cualitativa recolectada.

$$P(A/B) = \frac{P(A)P(B/A)}{P(B)} \quad (1)$$

La clasificación obtenida a partir de la ecuación 1 permite asignar un sentido al texto estudiado para poder clasificarlo de la manera más aproximada. De esta manera, se puede realizar inferencia para comprender la información textual para destacar las características y sentido del autor.

Información utilizada

En el Proyecto “Fortalecimiento del Sistema de C&CTI del Valle del Cauca: hacia una Economía del conocimiento”, con el fin de tener una medición y evaluación aproximada de la política de C&CTI para el Valle del Cauca, se llevaron a cabo 31 entrevistas semiestructuradas a diversos actores representativos del Sistema de Competitividad, Ciencia, Tecnología e Innovación para conocer se percepción sobre el estado actual del Sistema (ver Capítulo 1).

³ La regla de Bayes calcula la probabilidad de que se cumpla la condición de pertenecer a A dado B de la siguiente forma: multiplica la probabilidad de A y la de B dado A y la divide por la probabilidad de B. De esta forma estima la probabilidad condicionada, este método se puede extender a más de una categoría.

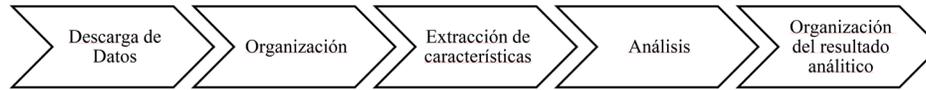


Figura 2.1 Proceso de análisis de datos textuales

Fuente: adaptado de Kwartler (2017).

De las 28 preguntas realizadas en total (en el Anexo 1 se presenta la guía utilizada para las entrevistas), 7 de estas fueron las seleccionadas para este análisis de sentimientos. Esta selección se realizó con base en secciones clave de la entrevista como: estado actual del Sistema de C&CTI; fortalezas y debilidades del Sistema; actores relevantes en la dinámica del Sistema; competitividad; y desafíos, retos y recomendaciones. Específicamente, las preguntas analizadas son:

- ¿Cómo describe usted el estado actual del Sistema de Competitividad, Ciencia, Tecnología e Innovación del Valle del Cauca?
- ¿Cuáles diría usted que son las principales fortalezas del Sistema de Competitividad, Ciencia, Tecnología e Innovación del Valle del Cauca?
- ¿Qué debilidades identifica usted? ¿qué debilidades se pueden resaltar del Sistema de Competitividad, Ciencia, Tecnología e Innovación?
- ¿Quiénes son los actores que hoy en día son relevantes en esta dinámica de Competitividad, Ciencia, Tecnología e Innovación del Valle del Cauca?
- ¿Cómo describes la competitividad en del departamento del Valle del Cauca? ¿Qué mecanismos son necesarios para el aumento de la competitividad en el departamento?
- En términos de retos, desafíos y recomendaciones ¿Cuáles serían los principales retos y desafíos del sistema en términos de recursos, políticas públicas, relaciones interinstitucionales, de sistemas de seguimiento, de aliados, de condiciones habilitantes?
- ¿Cuáles serían las 3 recomendaciones que priorizaría para el SC&CTI del Valle del Cauca?

Asimismo, estas preguntas fueron seleccionadas al ser lo suficientemente amplias para dar una buena percepción del estado actual del Sistema de C&CTI, siendo las más contestadas en la encuesta. Es importante resaltar que existen encuestas en las que no se respondieron todas las preguntas. De las preguntas elegidas existe un porcentaje de vacíos del 22% al 68%. La Tabla 2.1 muestra el número de respuestas vacías por pregunta y el porcentaje de vacíos con respecto al total de las encuestas.

Tabla 2.1 Porcentaje de vacíos en las preguntas elegidas

Pregunta	Vacíos	Porcentaje de vacíos
¿Cómo describe usted el estado actual del Sistema de Competitividad, Ciencia, Tecnología e Innovación del Valle del Cauca?	7	22.6%
¿Cuáles diría usted que son las principales fortalezas del Sistema de Competitividad, Ciencia, Tecnología e Innovación del Valle del Cauca?	9	29.0%
¿Qué debilidades identifica usted? ¿qué debilidades se pueden resaltar del Sistema de Competitividad, Ciencia, Tecnología e Innovación?	13	41.9%
¿Quiénes son los actores que hoy en día son relevantes en esta dinámica Competitividad, Ciencia, Tecnología e Innovación del Valle del Cauca?	12	38.7%

Pregunta	Vacíos	Porcentaje de vacíos
¿Cómo describes la competitividad en del departamento del valle del cauca? ¿Qué mecanismos son necesarios para el aumento de la competitividad en el departamento?	10	32.3%
¿Cuáles serían los principales retos y desafíos del sistema en términos de recursos, políticas públicas, relaciones interinstitucionales, de sistemas de seguimiento, de aliados, de condiciones habilitantes?	21	67.7%
¿Cuáles serían las 3 recomendaciones que priorizaría para el SC&CTI del Valle del Cauca?	11	35.5%

Fuente: elaboración propia.

Resultados

Análisis textual del SC&CTI en el Valle del Cauca

Como se mencionó en la sección de metodología a cada pregunta se le extraerá el *Wordcloud*, los sentimientos más dominantes por pregunta y en caso de ser relevante se caracterizarán las palabras más utilizadas de acuerdo al nivel educativo de los expertos. Los resultados se presentan para cada una de las preguntas seleccionadas.

¿Cómo describe usted el estado actual del Sistema de Competitividad, Ciencia, Tecnología e Innovación en el Departamento?

Esta pregunta busca establecer el estado actual del Sistema de C&CTI en el Departamento y la opinión de los expertos sobre este. La Figura 2.2 presenta la nube de palabras de esta pregunta.

En esta pregunta se resalta entre las palabras más importantes después de sistema, y la ciencia, tecnología, innovación y competitividad, expresiones como investigación, empresas, actores, universidades, proyectos, desarrollo, recursos, entre otros.

Es posible establecer que el Sistema de C&CTI es percibido por los expertos bajo un entorno positivo.

Así, frente al estado actual del Sistema de C&CTI, se destaca el papel que juegan, justamente, el desarrollo de la ciencia y la innovación, como piedra angular en la creación y aplicación de conocimiento, el aumento y sostenibilidad de la competitividad, el desarrollo y financiación de proyectos. Entre los actores que se destacan están las universidades y las empresas.

Por su parte, las expresiones usadas por los expertos difieren un poco de acuerdo a si tienen o no estudios de pregrado universitario. La Figura 2.3 muestra un diagrama de barras sobre las palabras usadas por estos dos tipos de expertos. Los primeros mencionan la participación de las empresas y la competitividad como elementos relevantes, mientras que el segundo grupo señala la importancia de las ciudades y el conocimiento. Este análisis resalta la heterogeneidad de las opiniones presentadas y las estrategias para incentivar la innovación deben ser también acordes a estas diferencias.

¿Cuáles diría usted que son las principales fortalezas del Sistema de Competitividad, Ciencia, Tecnología e Innovación del Valle del Cauca?

Esta pregunta busca establecer los elementos regionales que han sido fortalezas del Sistema de C&CTI. El *Wordcloud* (ver Figura 2.4) muestra entre las fortalezas el sistema, los procesos, la universidad, las entidades como la cámara de comercio, empresas y la educación. Sin embargo, se puede resaltar que existe una heterogeneidad en muchos mensajes que hablan de la gobernación, recursos, dinámica, capacidades, el talento, la formación y el conocimiento.

En línea con estos resultados, en el análisis cualitativo general de las entrevistas, se encontró que, para los entrevistados, una de las principales fortalezas del Sistema es el capital humano, es decir, la experiencia y el conocimiento de los actores del Sistema de C&CTI.



Figura 2.4 Nube de palabras - Fortalezas del SC&CTI

Fuente: elaboración propia con base en información de las encuestas semiestructuradas realizadas a los principales actores involucrados en el Sistema de C&CTI del Departamento del Valle del Cauca.

El análisis de las palabras más utilizadas, diferenciando entre expertos con pregrado universitario y expertos sin dicho nivel educativo, muestran nuevamente diferencias entre las fortalezas percibidas. La Figura 2.5 muestra que las palabras más utilizadas para los expertos que no tienen pregrado son: universidad, educación, y competitividad. Estas tres palabras evidencian la importancia dada a los procesos educativos por parte de los expertos. Las palabras clave para expertos con pregrado universitario son sistema, empresas, innovación y comercio. Esto evidencia la importancia dada al componente empresarial como entorno clave para el desarrollo y la consolidación del Sistema C&CTI en el Valle del Cauca.

¿Qué debilidades identifica usted? ¿qué debilidades se pueden resaltar del Sistema de Competitividad, Ciencia, Tecnología e Innovación?

Esta pregunta permite a los expertos hablar sobre las debilidades en el Sistema C&CTI del Valle del Cauca. La Figura 2.6 muestra el *Wordcloud* donde se resaltan el acceso y los recursos. Estas palabras pueden referirse a dificultad de acceso y falta

de recursos para el diseño y ejecución de diferentes propuestas/proyectos de innovación o tecnología para mejorar los procesos productivos de las empresas de la región. Asimismo, la palabra articulación se refiere, como expresaron algunos entrevistados, a los retos existentes relacionados con la generación de sinergias entre los diferentes actores del Sistema de C&CTI.

Las palabras más utilizadas, diferenciando entre expertos con pregrado universitario y expertos sin dicho nivel educativo, nuevamente muestran gran diferencia en la percepción (ver Figura 2.7). Las personas con pregrado universitario mencionan el acceso, los recursos, espacios, ministerio, las empresas y desarrollo. Por otra parte, los expertos sin pregrado universitario señalan los recursos como una debilidad importante y la pandemia como referencias importantes.

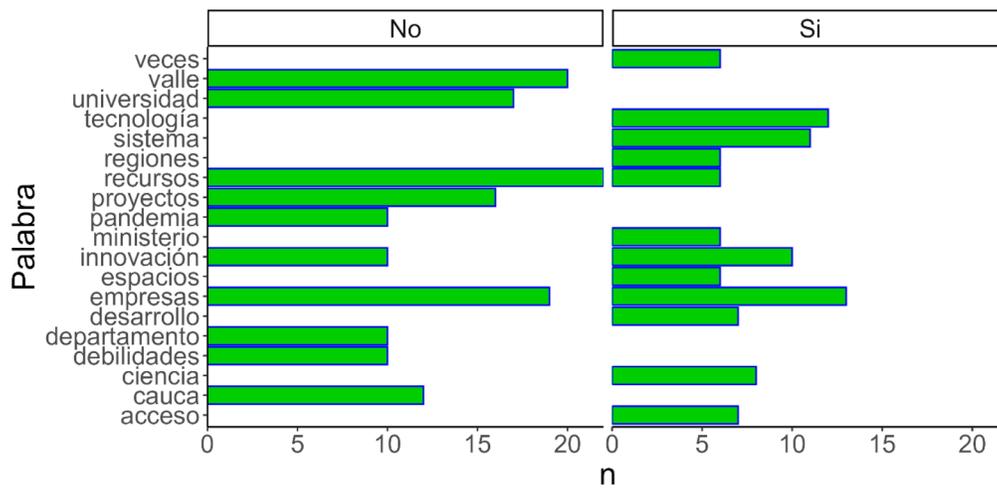


Figura 2.7 Uso de palabras: expertos con pregrado universitario vs expertos sin pregrado universitario – Debilidades del SC&CTI

Fuente: elaboración propia con base en información de las encuestas semiestructuradas realizadas a los principales actores involucrados en el Sistema de C&CTI del Departamento del Valle del Cauca.

¿Quiénes son los actores que hoy en día son relevantes en esta dinámica de Competitividad, Ciencia, Tecnología e Innovación del Valle del Cauca?

Esta pregunta permite identificar a los actores relevantes en la dinámica de C&CTI del Valle del Cauca. La Figura 2.8 muestra el *Wordcloud* de las respuestas a esta pregunta, donde se destacan como actores relevantes a las universidades, la cámara de comercio, las empresas, la Gobernación y el clúster.

En la Figura 2.9 se pueden identificar las diferencias en las respuestas a esta pregunta según si el experto tiene un pregrado universitario o no. Se observa que para los expertos sin pregrado la investigación, la educación, las universidades y la cámara de comercio son los actores relevantes. De igual forma, para los expertos con pregrado las universidades y cámara de comercio también son actores relevantes, adicionalmente, otros actores mencionados son la Gobernación y el clúster.

¿Cuáles serían los principales retos y desafíos del sistema en términos de recursos, políticas públicas, relaciones interinstitucionales, de sistemas de seguimiento, de aliados, de condiciones habilitantes?

Esta pregunta muestra los retos y desafíos del Sistema de C&CTI. En la Figura 2.10 se muestran las palabras más usadas en esta pregunta. Se resaltan los recursos, alianzas, credibilidad, software, desarrollo, proyectos y confianza.

En la Figura 2.11 se evidencian las respuestas más comunes de los expertos, diferenciando entre expertos con pregrado universitario y expertos sin dicho nivel educativo. Se observa que los retos para los expertos sin pregrado consisten en el software, la tecnología y el agregado que se puede generar. Mientras que los expertos con pregrado argumentan que el reto y desafío se concentra en recursos y credibilidad de las empresas.

¿Cuáles serían las 3 recomendaciones que priorizaría para el SC&CTI del Valle del Cauca?

Esta pregunta se concentra en las recomendaciones de los expertos y la Figura 2.12 muestra el *Wordcloud* de palabras más utilizadas. Se encuentran como elementos importantes de las respuestas: las capacidades, los proyectos, las universidades, los sectores, la cámara de comercio, los recursos y la articulación. Estas logran definir un amplio espectro de mejoras a tener en cuenta para priorizar en el desarrollo.

Las respuestas a esta pregunta no presentan diferencias en las recomendaciones que puedan evidenciar experiencias diferentes entre los expertos consultados.

Análisis de Conglomerados del SC&CTI del Valle del Cauca

Ahora bien, dentro del análisis de palabras se realizó un diagrama de análisis de conglomerados, el cual es un recurso de grupos, nodos o casos de tema que comparten palabras similares, valores de atributo o codificación. El diagrama de mapa ramificado o dendograma permite la clasificación de clústeres por categorías que comparten palabras similares. Para esto, a través software Nvivo 12, se aplicó una codificación automática de temas, donde el software identificó las palabras clave y aquellas palabras secundarias asociadas a estas y que presentaban una mayor frecuencia dentro de todas las entrevistas realizadas.

Seguidamente, se realizó un análisis de conglomerados alrededor de las diferentes palabras centrales y secundarias, de tal manera que se organizaron en clústeres de acuerdo a la similitud y cercanía de las palabras, en la forma que se presentaban en la verbalización de los entrevistados. La Figura 2.13 presenta los resultados del dendograma.

El primer clúster de palabras, aquellas de la parte superior conformadas por los colores azul y verde donde se encuentran "inversión" "puerto" "aeropuerto" "Buenaventura" y "Colombia", se interpreta principalmente desde los aspectos estructurales o físicos que caracterizan al departamento del Valle del Cauca. En el discurso de los entrevistados, la inversión

en el departamento se relaciona principalmente con las infraestructuras portuarias y áreas.

El segundo clúster de palabras, aquellas de color café donde se encuentran "gente" "local" "departamental" "gobierno" "nacional" "haciendo", estaría relacionado con categorías territoriales, y actividad del Gobierno tanto al nivel local, regional y nacional. En el discurso de los entrevistados fue recurrente estas menciones al hablar del Sistema de C&CTI, tanto a nivel local, comprendido por los diferentes municipios del Valle del Cauca, como a nivel departamental y nacional, donde hubo opiniones y percepciones sobre el Sistema Nacional de Ciencia, Tecnología e Innovación, y a su vez, diferentes narrativas de comparación entre el funcionamiento del Sistema de C&CTI en el Valle del Cauca, Antioquia y Cundinamarca.

Un tercer clúster de palabras, el cual presenta mayor tamaño frente a los otros, donde se encuentran aquellas palabras de color morado, habla principalmente de los diferentes recursos asociados al Sistema de C&CTI, donde las menciones se relacionan con "tecnología", "internet", "señal", pero también con el sector productivo con palabras como "comercio", "empresas", "puertos" y "colegios". Desde este clúster, las narrativas de los actores estarían describiendo principalmente actores, sectores y herramientas clave para el desarrollo del Sistema de C&CTI del Valle del Cauca, en el marco de procesos que lo lleven a escenarios de mayor competitividad.

El siguiente clúster, integrado por las palabras de color azul claro, se encuentra relacionado con las menciones de los actores frente a las capacidades de diseñar y ejecutar proyectos en ciencia, tecnología e innovación. Por lo tanto, en las narrativas de los entrevistados, palabras como "proyectos" "sector" "jalonar" "hacer" "empresarios" "financiación" describen el quehacer de los actores del Sistema de C&CTI en relación con sus capacidades de producción, pero también a las necesidades de financiación de los proyectos.

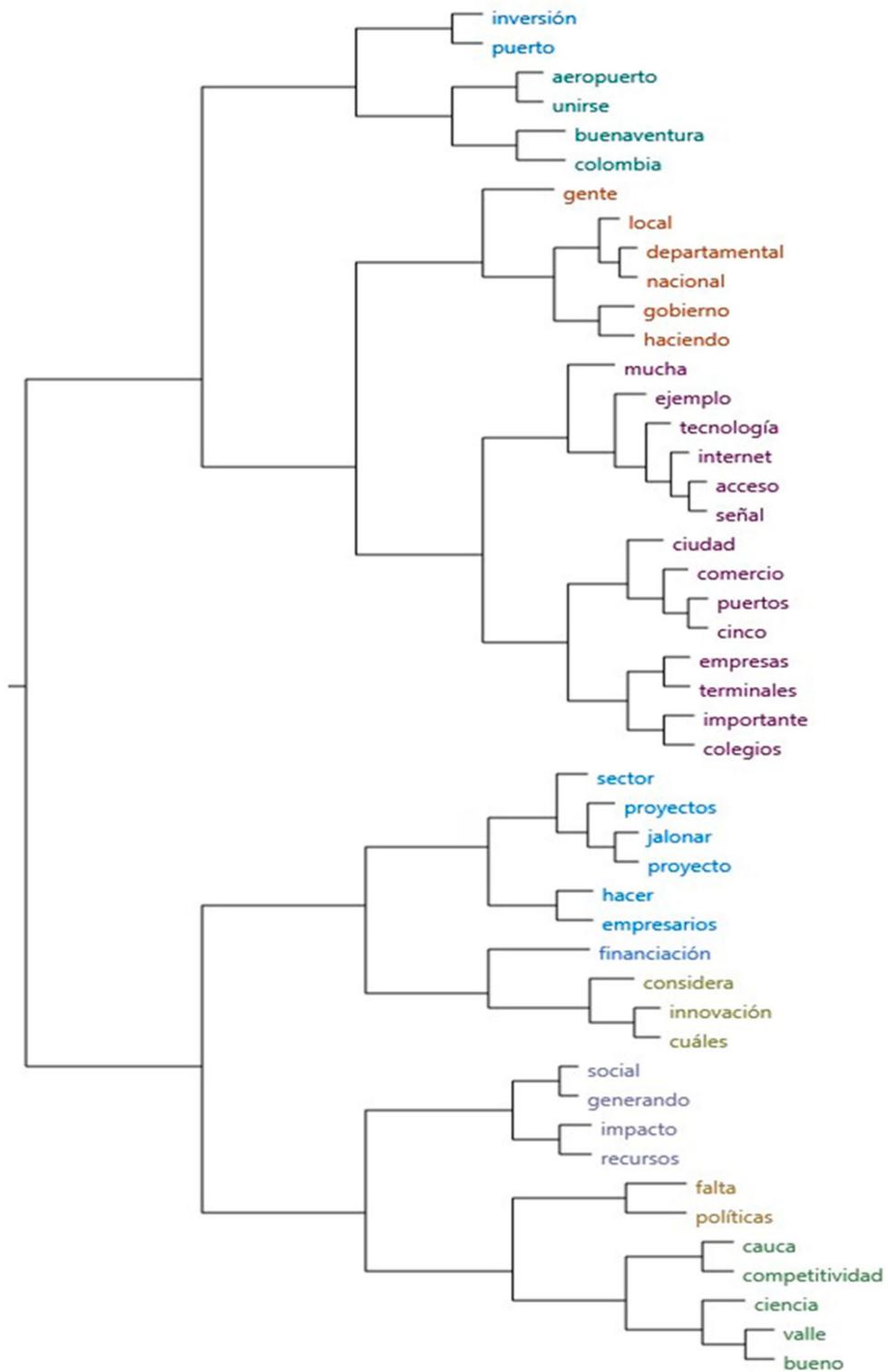


Figura 2.13 Análisis de conglomerados codificación temática

Fuente: elaboración propia.

Este análisis también permite establecer las diversas aproximaciones para incentivar y aprovechar las fortalezas planteadas por los mismos grupos. La mayoría de los expertos señala la importancia de las empresas y las universidades como actores que deben interactuar y de ser posible profundizar las relaciones que tienen. Es vital establecer un acceso claro a recursos y permitir que las instituciones como la cámara de comercio logren impulsar estas políticas.