

Uso de la minería de texto y revisión sistémica en R sobre el Patrimonio biocultural y su valorización económica.

Araceli Esmeralda Villarreal Flores

Naim Manríquez García

Sergio Colín Castillo

El patrimonio biocultural constituye un concepto integrador que enlaza la diversidad biológica y cultural, pero su valoración económica sigue siendo incipiente en contextos vitivinícolas como Parras, Coahuila. Para abordar este vacío, realizamos una revisión bibliográfica exhaustiva orientada por el método SALSA (Search–Appraisal–Synthesis–Analysis), complementada con minería de texto en R para profundizar en los patrones léxicos de la literatura disponible.

La búsqueda se realizó de manera manual en bases académicas (Scopus, Google Scholar, Crossref y Semantic Scholar) y literatura gris, consolidando 150 referencias iniciales, organizadas con Mendeley. A partir de este conjunto, se aplicó un riguroso proceso de selección y evaluación. Se establecieron criterios de depuración basados en el tipo de fuente, la existencia de revisión por pares, la claridad metodológica y la relevancia temática para la investigación. Este filtro permitió conformar una muestra final de 23 documentos de alta calidad, que constituyeron el corpus central para el análisis.

Con este corpus, la estrategia metodológica combinó la síntesis bibliográfica con el análisis computacional de palabras. En R, y mediante los paquetes *tm* y *tidytext*, se limpiaron los textos (eliminación de stopwords, estandarización de términos) y se calcularon frecuencias relativas. Las nubes de palabras y gráficos de barras evidenciaron la centralidad de conceptos como *biocultural*, *cultural*, *diversity*, *heritage*, *tourism*, *development*, *wine* y *landscape*, subrayando la estrecha relación entre patrimonio, identidad territorial y enoturismo.

Posteriormente, se exploraron coocurrencias de términos con *widyr*, y se construyeron grafos mediante *igraph* y *ggraph*. La aplicación del algoritmo de Louvain permitió identificar comunidades semánticas: en el bloque de gobernanza, por ejemplo, emergieron agrupaciones

alrededor de *rights*, *management* y *systems*; mientras que en el bloque de enoturismo se asociaron *wine*, *landscape* e *identity*. Esta aproximación estructural mostró que la literatura converge en torno a la conservación biocultural y el desarrollo turístico, pero también reveló vacíos sobre la distribución de beneficios económicos y la sostenibilidad social.

En paralelo, se diseñó una fase de análisis de sentimientos, que permitirá identificar la carga afectiva de los textos. Se espera que los sentimientos negativos se vinculen a marcos institucionales restrictivos, y los positivos a experiencias colaborativas en enoturismo. Esta perspectiva complementa el análisis cuantitativo, al capturar no solo la presencia de términos, sino también las emociones que orientan los debates académicos.

El aporte central de este trabajo radica en mostrar cómo la minería de texto en R, aplicada a un corpus académico cuidadosamente filtrado, constituye una herramienta potente para mapear tendencias, detectar vacíos y organizar discusiones sobre la valorización del patrimonio biocultural. Para la academia, ofrece un panorama reproducible que conecta modelos de valoración económica, gobernanza e identidad territorial. Para actores locales, sugiere que cualquier estrategia económica debe integrar tradiciones y recursos naturales en proyectos sostenibles. Finalmente, para la comunidad de LatinR, demuestra que la incorporación rigurosa de análisis textuales computacionales en ciencias sociales no solo es posible, sino necesaria para abrir diálogos interdisciplinarios.