Chapter 7 / Capítulo 7

Applied bibliometrics. From data to publication (English Edition)
ISBN: 978-9915-9680-7-0
DOI: 10.62486/978-9915-9680-7-0.ch07

Heat Maps: Representing Research Density / Mapas de Calor: Representando la Densidad de la Investigación

7.1. Interpretation of heat maps in bibliometric analysis

Heat maps are a powerful visual tool that transforms multidimensional bibliometric data into intuitive representations where color intensity encodes the density, frequency, or impact of research activity. Unlike networks that emphasize structural relationships, heat maps capture patterns of thematic, geographic, or temporal concentration through color gradients, enabling rapid identification of areas of high productivity, specialized niches, and knowledge gaps. Proper interpretation of these maps requires understanding that each shade represents a continuous value, where warm colors (reds, oranges) typically indicate high density or frequency. In contrast, cool colors (blues, greens) indicate areas of lower research intensity.

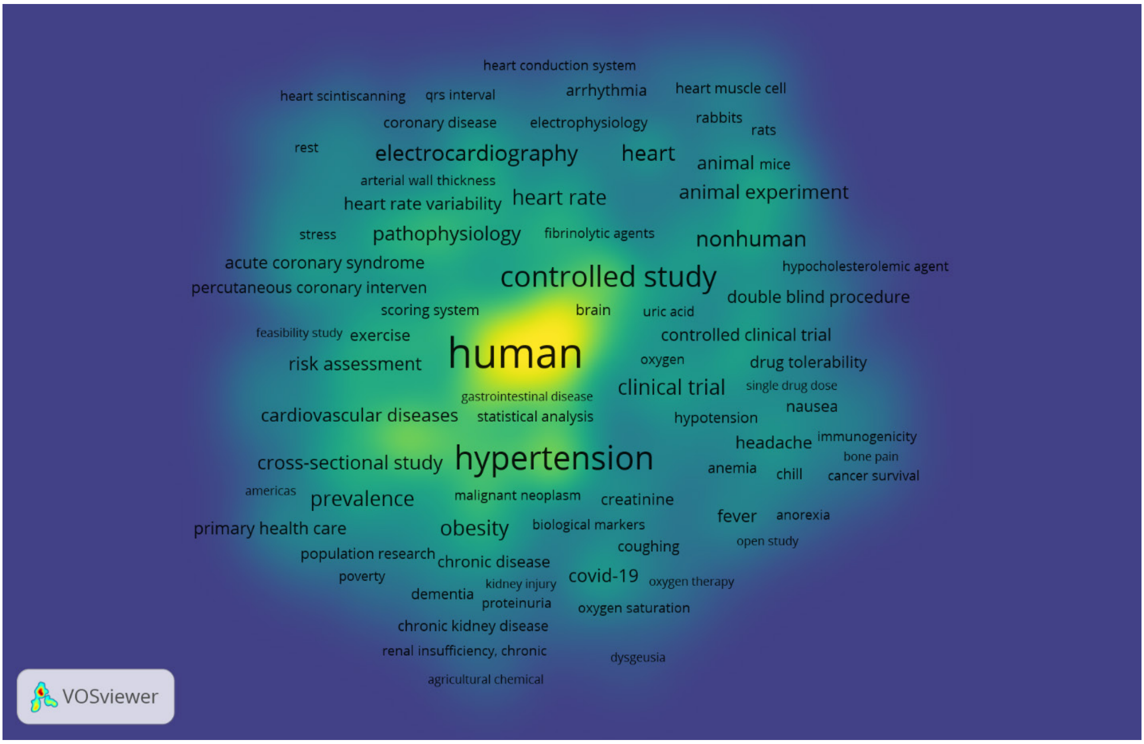

Figure 7.1. Example of a density map

Stratified reading of a bibliometric heat map involves breaking the visualization into multiple layers of meaning. The first layer, purely quantitative, reveals basic frequency distributions, where publications are concentrated, which topics receive the most attention, and which institutions lead production. The second layer, temporal, shows diachronic evolution when the map incorporates a chronological dimension, allowing us to track the migration of research interests, the emergence of new lines, and the decline of established paradigms. The third layer, relational, emerges when the heat map is superimposed with structural networks, revealing how high-density areas relate to central nodes and interdisciplinary bridges.

Contextualized interpretation transcends immediate visual analysis to integrate socio-institutional, epistemological, and science-policy factors that explain the observed patterns. An area of high research density may reflect both the intellectual vigor of a promising field and the effect of concentrated funding or passing academic fads. Cold spots may indicate unexplored frontiers with innovative potential, intellectual dead ends, or areas with methodological barriers to entry. The expert analyst distinguishes these possibilities by triangulating with other sources of evidence and specialized disciplinary knowledge, transforming the heat map from a mere distributional description into a diagnostic tool for strategic research planning.

7.2. Algorithm for constructing bibliometric heat maps

The methodologically rigorous construction of bibliometric heat maps begins with the precise definition of the dimensions to be represented, typically thematic, temporal, geographical, or institutional axes, and the selection of the intensity metric appropriate for the research objective. Metric options range from fundamental productivity indicators (e.g., number of publications) to sophisticated measures of impact (e.g., field-normalized citations) and specialization (e.g., thematic concentration indices). The choice of this metric fundamentally determines the type of patterns that the map will reveal, requiring careful alignment with the research questions that motivate the analysis.

Data processing for heat map construction involves successive stages of aggregation, normalization, and smoothing. Aggregation transforms individual publication data into summarized values for the cells of the multidimensional matrix that will constitute the map. Normalization adjusts these raw values to enable meaningful comparisons across domains with different sizes, publication practices, or citation traditions, which is essential for analyzing interdisciplinary fields. Smoothing applies interpolation algorithms to create gradual transitions between adjacent cells, improving visual readability but introducing potential artifacts that the analyst must recognize and control.

Color coding represents the stage where numerical values are transformed into interpretable visual experiences. The selection of color palettes must consider principles of visual perception, accessibility for people with color blindness, and established disciplinary conventions. Sequential palettes, with variations in brightness of a single color, are ideal for representing unidirectional magnitudes, while divergent palettes, with two contrasting colors, adequately capture deviations from a central reference point.

The choice of breakpoints between color intervals can highlight or hide critical patterns, requiring explicit methodological justification based on natural statistical distributions of the data or substantive thresholds significant to the field of study.

Interpretive validation closes the construction process, ensuring that emerging visual patterns correspond to real phenomena in the research ecosystem rather than methodological artifacts. This validation involves sensitivity testing of technical decisions (thresholds, smoothing algorithms, palettes), triangulation with other representations (networks, time series), and contrast with expert domain knowledge. The mature heat map thus transcends its initial descriptive function to become an interactive interface for analytical exploration. This dynamic tool allows hypotheses about the structure and evolution of scientific knowledge to be formulated and verified through the systematic manipulation of visualization parameters and the iterative exploration of different scales of analysis.

7.3. Creation of bibliometric heat maps

The creation of density maps in VOSviewer is integrated with graph generation, requiring only a change in the visualization in the “Items” tab to the “Density Visualization” option. This transition transforms the network representation into a heat map, where areas of intense color indicate regions of high element concentration, preserving the same spatial arrangement as in the network analysis. Density is calculated using essential functions that smooth the node point distribution, creating continuous gradients in which the color, from blue (low density) to red (high density), reveals thematic or collaborative clusters. Users can adjust the map's sensitivity using the smoothing parameter and customize the color palette to optimize readability based on the specific characteristics of the analyzed dataset.

In CitNetExplorer, density maps are generated using the “Cluster Density Visualization” function, which enables users to visualize the concentration of publications across different regions of the citation network. The process involves calculating the density of connections around each node and representing it in a heat map, with areas of greater research activity appearing in warm tones. This approach is particularly valuable for identifying periods of intense citation activity within specific research trajectories, revealing paradigmatic moments in the evolution of scientific fields.

CiteSpace implements density maps through its “Spectral Density Mapping” module, which combines spectral detection algorithms with thermal representations. The user must activate the “Show Density” option in the display control panel, which generates an additional layer of color that overlaps the conventional network map. This tool allows the identification of clusters with high internal cohesion and boundaries between different schools of thought, with the particularity that density is calculated by simultaneously considering spatial proximity and thematic similarity between elements.

In the R ecosystem, density maps are created using the *termDensity()* function for thematic analysis or *the authorDensity()* function for collaboration studies. The process first requires generating the co-occurrence or collaboration matrix and then applying the *heatmap()* function to the normalized matrix. Advanced customization includes adjusting smoothing parameters using Gaussian kernels and defining specific color palettes for different density ranges. This programming-based approach offers maximum analytical flexibility but requires technical skills in matrix manipulation and R visualization.

PyBibX in Python provides density maps via its density_analysis module, which implements kernel density estimation algorithms for bibliometric distributions. The typical workflow involves loading the bibliographic dataset, calculating multidimensional coordinates through dimensionality reduction, and then generating the heatmap with *plot_density_map()*. The library allows you to adjust the kernel bandwidth to control the level of smoothing and export maps in vector formats for high-quality publications. This implementation is compelling for analyzing large volumes of data, where efficient processing and advanced customization of visualization are required.

Recap

- Heat maps are graphical representations that show the intensity or density of a phenomenon using color gradients.
- In bibliometrics, they allow you to visualize thematic areas with varying concentrations of publications or citations.
- They are based on the distribution of frequencies or correlations within matrices of co-occurrence, co-citation, or collaboration.
- Each cell or point on the map reflects the level of research activity, measured by the number of documents, citations, or links between terms.
- Color acts as a visual variable: warmer tones (reds, oranges) indicate greater

density, and cooler tones (blues, greens) indicate less activity.

- Heat maps help detect emerging, consolidating, or declining areas within a scientific field.
- They also allow us to observe the interrelationship between topics or authors, revealing clusters of high research concentration.
- They are used in both thematic analysis (keywords) and collaboration analysis (authors, countries, institutions).
- The input data is obtained from databases such as Scopus, Web of Science, or Dimensions, and exported in a compatible format (CSV, RIS, BibTeX).
- Creating a heat map requires a numerical matrix of similarities or frequencies.
- The process includes data normalization, choice of color scale, and configuration of the density range.
- The most commonly used tools for creating bibliometric heat maps are VOSviewer, Bibliometrix (R/Biblioshiny), CiteSpace, Gephi, and Excel with statistical add-ons.
- In VOSviewer, density maps represent areas of higher element concentration (nodes) with more intense colors.
- Thematic density maps facilitate the identification of conceptual nuclei and relationships between areas of knowledge.
- Institutional or geographic maps show the spatial distribution of scientific production and international collaboration.
- These maps allow comparison of scientific performance across regions or disciplinary fields.
- The use of appropriate scales avoids visual distortions and improves the interpretation of data density.
- Heat maps should be accompanied by clear legends and complete metadata explaining the color and scale criteria.
- A rigorous interpretation combines visual reading with complementary statistical analysis (frequencies, correlations, centralities).
- When used correctly, heat maps become a powerful tool for communicating bibliometric results in a visual, intuitive, and comparative manner.

Self-assessment questions

1. What do the colors represent in a heat map applied to bibliometrics?
2. What type of quantitative information is displayed in a bibliometric heat map?
3. What is the difference between a thematic heat map and an institutional heat map?
4. What databases are typically used to generate the input data?
5. What role does standardization play in map creation?
6. What programs allow you to create heat maps of research density?
7. How does VOSviewer interpret the hottest areas on the map?
8. What precautions should be taken when choosing the color scale?
9. Why is it essential to include an explanatory legend and metadata?
10. How does the use of heat maps contribute to the visual understanding of scientific output?

BIBLIOGRAPHY

1. Chen C. Mapping scientific frontiers: The quest for knowledge visualization. 2nd ed. London: Springer; 2017. doi: 10.1007/978-3-319-32043-4.

2. Börner K. Atlas of science: Visualizing what we know. Cambridge (MA): MIT Press; 2010. ISBN: 9780262014454.

3. Börner K, Polley DE. Visual perspectives: A practical guide to making sense of data. Cambridge (MA): MIT Press; 2014. ISBN: 9780262027898.

4. Van Eck NJ, Waltman L. Software survey: VOSviewer, a computer program for bibliometric mapping. Scientometrics. 2010;84(2):523–38. doi: 10.1007/s11192-009-0146-3.

5. Klavans R, Boyack KW. Toward a consensus map of science. J Assoc Inf Sci Technol. 2009;60(3):455–76. doi: 10.1002/asi.20991.