

The diffusion of NLP methods in marketing research: a systematic analysis

Olivier Caron, Christophe Benavent
DRM, Université Paris-Dauphine, Université PSL, CNRS, Paris, 75016, France

Introduction

Drawing on the Technology Acceptance Model (TAM), this research navigates the evolving publication landscape marked by increased competition and organized around journal rankings and citation scores. This dynamic forces researchers to:

- Adopt publication strategies
- Balance the pressure to publish, the quest for originality and the legitimacy of their methods
- Meet the academic community expectations through novelty

Methods

Systematic Literature Review

- Conducted a comprehensive search on Scopus database as of August 18, 2023 using the Scopus API with Rscopus R package
- Keywords: "NLP", "natural language processing", "text mining", "text analysis"
- Focused on articles in journals with "marketing" or "consumer" in the title.
- Manual filtering to exclude irrelevant articles.

Structural Topic Modeling (STM)

- STM to understand the evolution of research themes over time.
- The text used for topic analysis is a combination of the articles' titles, abstracts, and keywords.
- Used the year of publication as a covariate in the STM algorithm to analyze how topic prevalence changes over time.

Author Network Analysis

- Analyzed author collaborations from 1996 to 2015 and 1996 to 2023 using networkX and ipysigma modules in Python
- Used Louvain algorithm to identify communities within the author network.
- Node size is proportional to the number of citations received by the researcher's articles.
- Edge thickness represents the number of collaborations between researchers.
- Graph Layout used is ForceAtlas2
- The ten largest communities identified by the Louvain algorithm are color-coded.

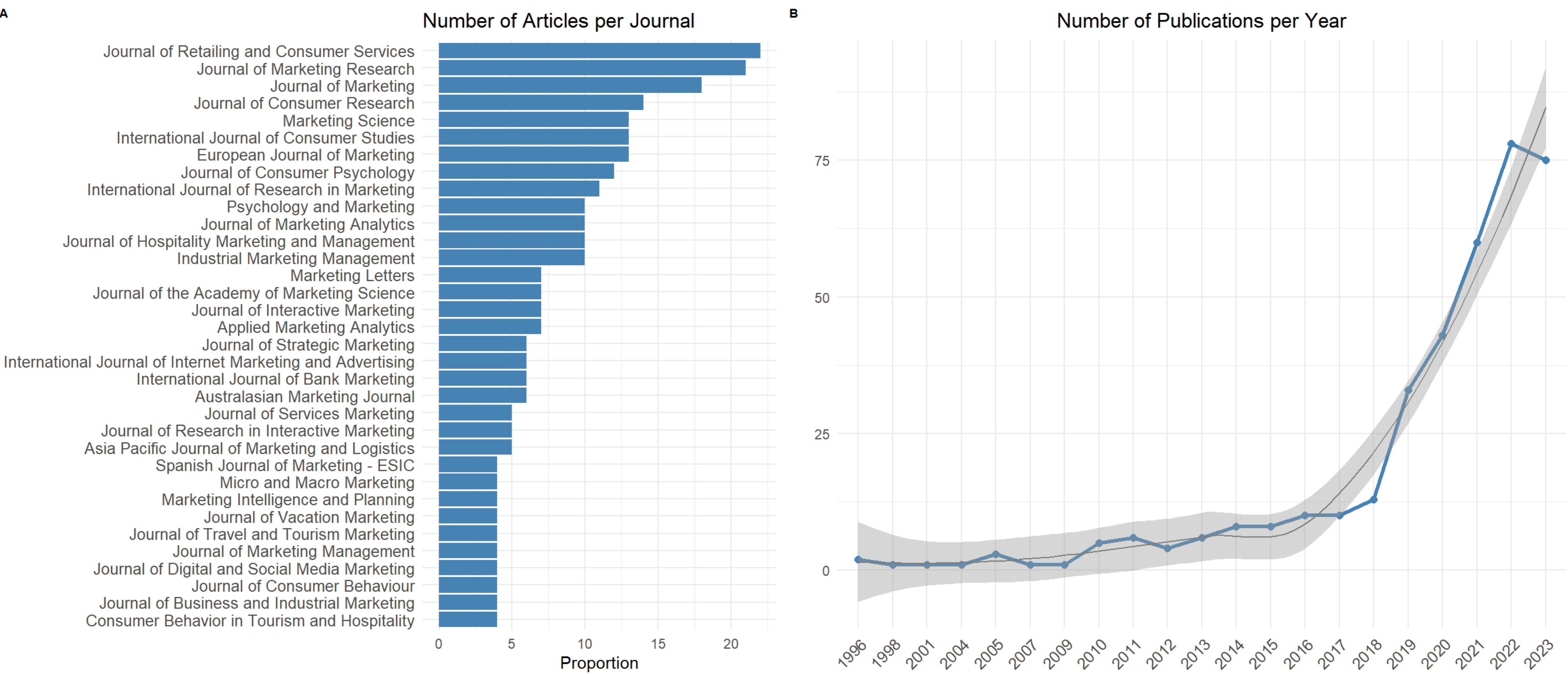
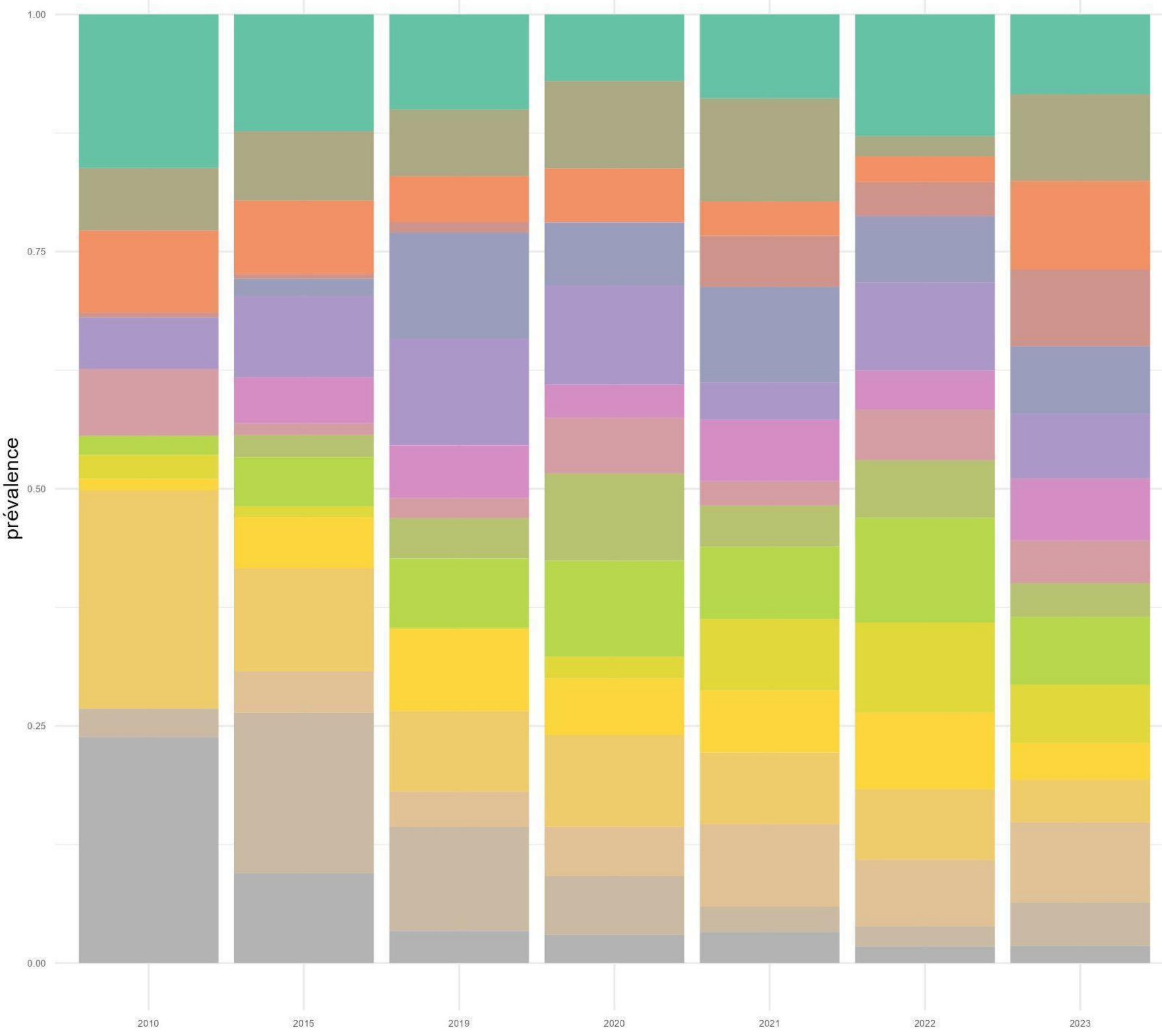
Results

NLP as an Academic Trend

- Significant increase in NLP-related publications in marketing after 2018
- NLP adoption was initiated in high-ranked journals seeking originality.
- These journals legitimize the methods, influencing second-tier publications as well as researchers

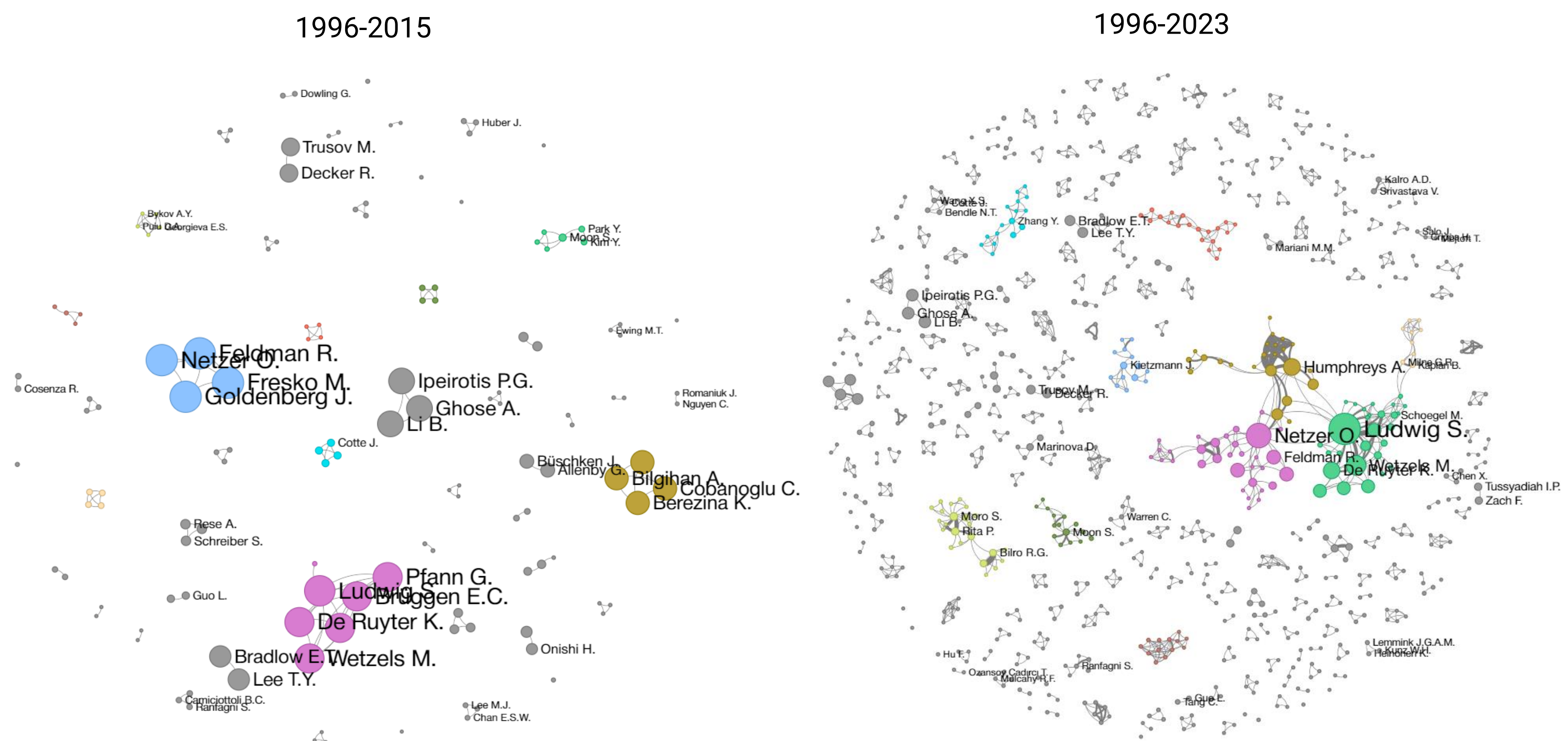
A concentric diffusion of topics

- The number of topics identified by STM increases steadily over time.
- Topics accumulate rather than substitute previous ones, indicating a broadening of research areas.
- Each significant increase in topics aligns with the emergence of new data sets and research problems.



Fragmented Growth in Author Collaborations

- Increased number of authors from 116 (1996-2015) to 923 (1996-2023).
- Collaboration density decreased from 1.84% to 0.32%.
- Major communities concentrate most citations.



Discussion

Adoption Dynamics

The adoption of NLP in marketing research reflects a significant shift driven by both technological advances and data availability. The diffusion process is mainly top-down, originating from high-impact publications. Researchers initially faced high learning costs, but these have decreased over time due to the formation of internal communities of practice within institutions and the availability of online tutorials and resources.

Localized Adoption

NLP methods are mainly adopted in specific research areas, driven by new questions, data availability, and suitable techniques. This creates a focused but growing use in the academic community.

Methodological Innovation

NLP methods are now well integrated into marketing research offering better data analysis and new methods. The integration of advanced NLP techniques, such as sentiment analysis and topic modeling, has enriched the methodological toolkit of marketing researchers, enabling more sophisticated analyses and deeper insights into consumer behavior and marketing strategies.

Conclusion

While the initial adoption of NLP methods was slow, it has now become an important approach in marketing research for it enables researchers to study large volume of qualitative data, driven by reduced learning costs and increased accessibility.

Future research will likely see further integration of advanced NLP techniques, facilitating more sophisticated analyses and deeper insights into consumer behavior and marketing strategies such as classification and sentiment analysis done with LLMs.

References

Abrahamson, E. (1991). Managerial Fads and Fashions: The Diffusion and Rejection of Innovations. *The Academy of Management Review*, 16(3), 586.

Davis, F. D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly*, 13(3), 319.

Roberts, ME, Stewart, BM, Tingley, D, Lucas, C, Leder-Luis, J., Gadarian, S. K., Albertson, B., & Rand, D. G. (2014). Structural Topic Models for Open-Ended Survey Responses. *American Journal of Political Science*, 58(4), 1064-1082.