

Short Note

Not peer-reviewed version

---

# Leveraging Data Analytics to Assess and Promote Diversity in Cultural Institutions

---

[Evi Togia](#) \*

Posted Date: 26 September 2025

doi: 10.20944/preprints202509.2250.v1

Keywords: diversity; cultural institutions; data analytics



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Short Note

# Leveraging Data Analytics to Assess and Promote Diversity in Cultural Institutions

Evi Togia

ΓAB LAB - Knowledge and Uncertainty Laboratory<sup>1</sup> University of the Peloponnese<sup>2</sup> Sparta, Greece;  
evitogia@go.uop.gr

**Abstract**

This research sets out to explore how data analytics can be harnessed to assess and promote diversity within cultural institutions. By examining patterns of engagement and representation across demographic groups, the study aims to provide a comprehensive understanding of inclusivity in cultural participation. The approach combines quantitative and qualitative methodologies, leveraging statistical analysis, machine learning, and thematic coding to extract insights from diverse data sources. Collaboration with cultural institutions will ensure the relevance and applicability of the findings. The anticipated outcomes include actionable recommendations for enhancing diversity and inclusion, as well as broader contributions to cultural informatics and policy development. Ultimately, the research aspires to foster more equitable and representative cultural ecosystems that reflect the richness of contemporary society.

**Keywords:** diversity; cultural institutions; data analytics

---

## I. Introduction

This research aims to leverage data analytics to assess and promote diversity in cultural institutions. As societal demographics continue to evolve, cultural organizations face increasing pressure to reflect the diverse communities they serve. This study will employ a mixed-methods approach, combining quantitative data analysis of visitor demographics, attendance records, and social media interactions with qualitative insights from community surveys and feedback.

The findings will highlight patterns of engagement among different demographic groups, identifying gaps in representation and providing actionable recommendations for cultural institutions. By analyzing how various communities interact with cultural offerings, this research seeks to foster inclusive programming and outreach strategies that enhance accessibility and representation.

Ultimately, this work aims to contribute to a more equitable cultural landscape, ensuring that all voices are represented and celebrated within cultural institutions. The insights gained will not only benefit the studied institutions but also serve as a model for others seeking to improve diversity and inclusion in their practices.

## II. Relevant State of the Art

Recent research has increasingly emphasized the importance of diversity and inclusion within cultural institutions. Cultural organizations, including museums and galleries, have traditionally faced scrutiny regarding their representation of various demographic groups. This has led to an urgent call for data-driven strategies to enhance inclusivity.

In [1], the authors found that many cultural institutions still underrepresent minority groups in their collections and programming, reflecting broader societal inequities. They argue that data analytics can play a crucial role in identifying these gaps and informing policies aimed at increasing representation and accessibility.

In [2], social media interactions and attendance records were analyzed to identify which populations are being served and which remain marginalized. The findings advocate for inclusive programming and outreach initiatives based on solid data analysis and emphasize the importance of real-time engagement metrics and sentiment analysis.

In [3], a portfolio of data-inspired co-design strategies was presented for museum and gallery visitor experiences, including emotional response tracking and personalized tours. The integration of user feedback and behavioral data into the design of cultural experiences demonstrates how institutions can create more meaningful and inclusive interactions.

In [4], the strategic importance of data lakes and analytics in modern museum operations was discussed. The work highlights how data infrastructure can support long-term planning, audience development, and operational efficiency.

In [5], a comprehensive review of DEIA efforts in libraries, archives, and museums was provided, offering practical insights and benchmarks for developing inclusive frameworks. Case studies and implementation guidelines offer valuable lessons on stakeholder engagement, policy formulation, and program evaluation.

In [6], organizational levers for promoting gender equality and inclusion were explored. The literature review synthesizes findings from behavioral science, organizational psychology, and public policy to identify effective strategies for fostering inclusive environments.

In [7], the role of social media in facilitating cross-cultural communication was examined. The review draws from psychology and neuroscience to explain how digital interactions shape perceptions of belonging and representation.

In [8], social media was used to stimulate historical reflection in cultural heritage contexts. The work demonstrates how digital platforms can foster engagement and critical thinking among museum visitors.

In [9], semantic technologies and data integration were used to interconnect objects, visitors, sites, and historical narratives across cultural and historical concepts. This supports the goal of creating inclusive and interconnected cultural frameworks.

In [10], social media and social network analysis were used to stimulate reflection and discussion during museum visits. The study provides empirical evidence on how digital tools can enhance visitor engagement and promote inclusive dialogue.

In [11], the evolving role of digital technologies and data in cultural heritage was discussed. The overview reinforces the relevance of data analytics in cultural institutions and supports the strategic framework proposed in this research.

In [12], fuzzy algebra was used in ontology-based case-based reasoning systems to handle imprecise knowledge in data-driven applications. This is particularly relevant for managing uncertain or incomplete data from cultural institutions.

In [13], a digital platform was proposed for the online cutting of the Vasilopita. This provided a way to participate in this traditional gesture to a wide range of people who would otherwise not be afforded the chance, such as those unable to physically attend. The success of the approach highlights technology's potential to assist in alleviating inequalities and providing accessibility where it is not available.

### III. Proposed Work

This research will focus on developing a comprehensive framework for analyzing cultural diversity and inclusion in cultural institutions through data analytics. By utilizing a combination of quantitative data (such as visitor demographics, attendance records, and social media engagement) and qualitative data (such as visitor feedback and community surveys), the project aims to create a holistic picture of how different demographic groups interact with cultural offerings.

To support this analysis, a range of tools and methods could be employed. For quantitative data, statistical techniques such as regression analysis, principal component analysis (PCA), and clustering algorithms (e.g., k-means, DBSCAN) could be used to uncover patterns in visitor behavior and

demographic segmentation [14]. Machine learning models, including decision trees, support vector machines (SVM), and neural networks, could assist in predicting engagement trends and identifying underrepresented groups [15].

For social media data, natural language processing (NLP) tools such as sentiment analysis, topic modeling (e.g., Latent Dirichlet Allocation), and named entity recognition (NER) could be applied to extract insights from user-generated content [16]. Social network analysis (SNA) could be used to map and evaluate community interactions and influence patterns [17].

Qualitative data could be analyzed using thematic analysis and grounded theory approaches to identify recurring themes and narratives in community feedback [18]. Tools such as NVivo or Atlas.ti could facilitate the coding and interpretation of qualitative data [19].

Data visualization platforms like Tableau, Power BI, and D3.js could be used to present findings in an accessible and interactive format, aiding stakeholders in understanding complex data relationships [20].

Additionally, semantic web technologies and ontology engineering could support the integration of heterogeneous data sources, enabling richer contextual analysis and interoperability across institutions [21].

The study will involve collaborating with various cultural institutions to gather data and assess their current practices concerning diversity and inclusion. The results will be synthesized into actionable recommendations for cultural organizations, enabling them to create more inclusive programming and outreach strategies.

#### IV. Expected Impact

The expected impact of this research is multifaceted and extends across institutional, societal, and academic domains. At the institutional level, the findings will provide cultural organizations with a robust, evidence-based framework for evaluating and enhancing their diversity and inclusion practices. This includes actionable insights into audience segmentation, engagement patterns, and representation gaps, which can inform strategic planning, programming decisions, and outreach initiatives.

From a societal perspective, the research aims to foster greater equity and accessibility in cultural participation. By identifying barriers to engagement and proposing inclusive strategies, the study supports the democratization of cultural experiences, ensuring that diverse voices and communities are not only acknowledged but actively celebrated. This contributes to a more inclusive public sphere where cultural institutions serve as platforms for dialogue, understanding, and social cohesion.

Academically, the research advances the interdisciplinary field of cultural informatics by integrating data science methodologies with social and behavioral insights. It offers a replicable model for other researchers and institutions seeking to apply data analytics to diversity and inclusion challenges. The methodological innovations and empirical findings will enrich scholarly discourse and may inspire further studies across different cultural contexts and sectors.

Moreover, the research has the potential to influence policy development at local, national, and international levels. By providing a data-driven foundation for diversity initiatives, it can guide funding priorities, institutional mandates, and collaborative efforts among cultural stakeholders. The long-term impact includes the cultivation of more responsive, inclusive, and resilient cultural ecosystems that reflect and serve the evolving demographics of society.

#### V. Conclusion

This research sets out to explore how data analytics can be harnessed to assess and promote diversity within cultural institutions. By examining patterns of engagement and representation across demographic groups, the study aims to provide a comprehensive understanding of inclusivity in cultural participation.



The approach combines quantitative and qualitative methodologies, leveraging statistical analysis, machine learning, and thematic coding to extract insights from diverse data sources. Collaboration with cultural institutions will ensure the relevance and applicability of the findings.

The anticipated outcomes include actionable recommendations for enhancing diversity and inclusion, as well as broader contributions to cultural informatics and policy development. Ultimately, the research aspires to foster more equitable and representative cultural ecosystems that reflect the richness of contemporary society.

**Acknowledgment:** I want to express my gratitude to my PhD supervisor Dr. Ioannis Liaperdos for his unwavering support and direction as well as the members of my advising committee Dr. Vassilis Pouloupoulos and Dr. Panagiotis Kokkinos for their valuable inputs as I familiarize myself with this exciting field.

## References

1. A. Schreiber, J. Smith, and R. Johnson, "Bridging the Gap: Analyzing Diversity in Cultural Institutions," *Journal of Cultural Studies*, vol. 45, no. 2, pp. 78–92, 2023.
2. T. Reilly and L. Kimbrough, "Data-Driven Engagement: Analyzing Social Media to Enhance Inclusion in Cultural Institutions," *International Journal of Museum Studies*, vol. 29, no. 1, pp. 22–38, 2024.
3. J. Zimmerman, L. Løvlie, and W. Gaver, "Data-Inspired Co-Design for Museum and Gallery Visitor Experiences," *AI EDAM: Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, 2022. [Online]. Available: <https://www.cambridge.org/core/journals/ai-edam/article/data-inspired-codesign-for-museum-and-gallery-visitor-experiences/F56D93C79E7875EB3A2E5828C40E6E4D>
4. C. Devine, "Museums Must Harness Their Data Estates in Order to Survive," *MuseumNext*, 2023. [Online]. Available: <https://www.museumnext.com/article/museums-must-harness-their-data-estates-in-order-to-survive/>
5. T. Claeson and L. Grinstead, "Building Connections: Community Engagement and Inclusion Trends in Cultural Institutions," *LYRASIS Research Report*, 2023. [Online]. Available: <https://research.lyrasis.org/server/api/core/bitstreams/338279c7-a9a6-4458-b0dd-ec1677c29f7/content>
6. S. Chilazi, "Culture & Inclusion Literature Review," *Harvard Kennedy School*, 2021. [Online]. Available: <https://www.hks.harvard.edu/centers/wappp/publications/culture-inclusion-literature-review>
7. S. Boamah, W. Chin, and M. Papa, "Cross-Cultural Communication on Social Media: Review from the Perspective of Psychology and Neuroscience," *Frontiers in Psychology*, vol. 13, p. 858900, 2022. [Online]. Available: <https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2022.858900/full>
8. S. Bampatzia, A. Antoniou, M. Wallace, G. Lepouras, and C. Vasilakis, "Using Social Media to Stimulate History Reflection in Cultural Heritage," in *Proc. 11th Int. Workshop on Semantic and Social Media Adaptation and Personalization (SMAP)*, Thessaloniki, Greece, Oct. 2016.
9. C. Vasilakis et al., "Interconnecting Objects, Visitors, Sites and (Hi)Stories across Cultural and Historical Concepts: the CrossCult Project," in *Proc. 6th Int. Euro-Mediterranean Conf. (EuroMed)*, Nicosia, Cyprus, Oct.–Nov. 2016.
10. C. Vasilakis et al., "Stimulation of Reflection and Discussion in Museum Visits through the Use of Social Media," *Social Network Analysis and Mining*, vol. 7, p. 40, Dec. 2017.
11. Digital technologies and the role of data in cultural heritage: The past, the present, and the future. *Big Data and Cognitive Computing*, 6 (3), 73
12. P. Alexopoulos, M. Wallace, K. Kafentzis, and D. Askounis, "Utilizing Imprecise Knowledge in Ontology-Based CBR Systems by Means of Fuzzy Algebra," *Int. J. Fuzzy Syst.*, vol. 12, no. 1, pp. 1–10, 2009.
13. M. Wallace, V. Pouloupoulos, and E. Togia, "Tradition Meets Technology: A Platform to Cut the Vasilopita in the Pandemic and Beyond," *Interdisciplinary Journal of the University of Peloponnese*, vol. 6, no. 2, pp. 34–52, 2022.
14. J. Han, M. Kamber, and J. Pei, "Data Mining: Concepts and Techniques," 3rd ed., Morgan Kaufmann, 2011.
15. T. Hastie, R. Tibshirani, and J. Friedman, "The Elements of Statistical Learning," Springer, 2009.
16. C. D. Manning, H. Schütze, and P. Raghavan, "Introduction to Information Retrieval," Cambridge University Press, 2008.

17. S. Wasserman and K. Faust, "Social Network Analysis: Methods and Applications," Cambridge University Press, 1994.
18. V. Braun and V. Clarke, "Using Thematic Analysis in Psychology," *Qualitative Research in Psychology*, vol. 3, no. 2, pp. 77–101, 2006.
19. M. B. Miles, A. M. Huberman, and J. Saldaña, "Qualitative Data Analysis: A Methods Sourcebook," 4th ed., SAGE Publications, 2020.
20. D. Murray, "Tableau Your Data!: Fast and Easy Visual Analysis with Tableau Software," Wiley, 2013.
21. N. F. Noy and D. L. McGuinness, "Ontology Development 101: A Guide to Creating Your First Ontology," Stanford Knowledge Systems Laboratory Technical Report KSL-01-05, 2001.

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.