The Holistic Archival Personality Profiling Model (HAPPM): Comprehensive Data Integration for Personality Analysis

James Hutson
Pace Ellsworth

Follow this and additional works at: https://digitalcommons.lindenwood.edu/faculty-research-papers

Part of the Artificial Intelligence and Robotics Commons, and the History Commons
The Holistic Archival Personality Profiling Model (HAPPM): Comprehensive Data Integration for Personality Analysis

James Hutson\(^1\)*, Pace Ellsworth\(^2\)

\(^1\)Lindenwood University, Saint Charles, MO, United States of America
\(^2\)Preserving Records Everywhere (PRE), United States of America

\*Email: jhutson@lindenwood.edu

Abstract

The traditional approach to biographical profiling, predominantly reliant on limited and fragmented datasets, has frequently resulted in superficial personality understandings. This is largely due to an overemphasis on official records and notable events, neglecting the rich tapestry of everyday experiences and personal interactions that significantly shape personalities. To address this shortcoming, this article introduces a multi-disciplinary methodology, The Holistic Archival Personality Profiling Model (HAPPM), which integrates a diverse array of archival materials, including personal correspondences, social media footprints, and family memorabilia. This approach involves digitizing various data forms, including handwritten documents, into machine-readable text, and then semantically classifying this data with biotags, chronotags, and geotags for organization within specific spatial and temporal contexts. Such comprehensive data aggregation establishes a more accurate "space-time continuum" for individuals, enhancing our understanding of their lives. The innovative aspect of HAPPM is the utilization of large language models to "converse" with the data, facilitating a more holistic representation of personalities. Preliminary results from applying HAPPM have shown its efficacy in uncovering previously unknown aspects of individual lives, offering insights into personal beliefs, daily routines, and social interactions. This has been validated through comparative analysis with existing biographical data, revealing a more complete and nuanced understanding of personalities. Therefore, HAPPM marks a significant advancement in personality profiling, capturing not only the grandiose but also the mundane, and offering a comprehensive tool for researchers and historians to explore the full spectrum of human experience.

Keywords

Archival data aggregation, Personal profiling, Digital humanities, Contextualized history, Semantic analysis, Biographical reconstruction
Introduction

The realm of biographical profiling, a cornerstone in historical and psychological research, has been traditionally constrained by the limitations of its data sources, focusing primarily on official records and prominent events (Hadzigeorgiou, 2021; Smith, 2012). However, as highlighted by Neale, Henwood, and Holland (2012), this reliance on limited and fragmented datasets overlooks the essence of human experiences, particularly daily occurrences and personal interactions that shape personality. Consequently, the traditional method results in superficial portrayals, emphasizing recorded events and overlooking the everyday experiences that shape character and worldview (Clark & Watson, 1988; Dhruvarajan, 1981; Haehner et al., 2023; Reis, 2013). This bias towards 'significant' historical or societal events is evident in the profiles of historical figures like Abraham Lincoln or Marie Curie, where significant contributions are highlighted, but elements like personal letters and diaries that offer a more nuanced understanding are often missing (Jolly & Stanley, 2005; Pasachoff, 1996; Stanley, 2004; Thomas, 2008). Additionally, the survival of ephemera such as birthday cards or diaries is often contingent on the individual's perceived importance, leading to a narrative that may be incomplete or potentially misleading (Goodwin & Parsons, 2022; Hutnyk, 2023; Immler, 2023; Jolly, 2013; Melleville et al., 2022).

Addressing these shortcomings, the Holistic Archival Personality Profiling Model (HAPPM) introduces a paradigm shift in biographical profiling, integrating a broad spectrum of archival materials, including personal correspondences and social media interactions, to construct a more comprehensive and authentic representation of individuals (Xu et al., 2023). This approach involves digitizing diverse data forms into machine-readable text, allowing for semantic classification using biotags, chronotags, and geotags, and organizing it within specific spatial and temporal contexts (Ariano et al., 2017; Davoudi & Traviglia, 2023; Hu et al., 2022; Hussain et al., 2022; Tasse et al., 2017). A key innovation of HAPPM is the use of large language models (LLM) to analyze and synthesize this nuanced data, enabling a form of historical and biographical 'dialogue' that offers insights into the character, thought processes, and emotional states of individuals (Guntrip, 2018; Hutson & Ratican, 2023; Varshney et al., 2023; Wu et al., 2023). This methodology not only reconstructs a more comprehensive narrative of individual experiences but also addresses the limitations of traditional profiling, providing a dynamic "space-time continuum" for a deeper understanding of personalities (Kubler, 2008).

Material and Methods

The current study employs a multi-faceted approach in its exploration of personality profiling and historical biography, integrating a variety of methodologies each grounded in established academic literature and practices. This integration aims to offer a comprehensive and nuanced understanding of the subjects under study, addressing both the psychological intricacies of personality and the historical depth of individual biographies.

In personality profiling, traditional methods have predominantly relied on psychometric assessments and psychoanalytic theories, seeking to unravel the complexity of human character through structured tools and interpretative frameworks (Post, 2014). These approaches, ranging from the Freudian psychoanalytic theory to the Myers-Briggs Type Indicator, have provided foundational insights into the myriad facets of personality (Parker, 1998). Concurrently, the field
of historical biography has traditionally emphasized a narrative-driven, chronologically structured recounting lives of individuals, often prioritizing public achievements and documented events (Meister, 2018). This methodological approach has tended to focus on the impact of singular figures within the broader tapestry of historical events, thereby offering insights into their contributions and influence (Banner, 2009). Both domains, though distinct in their focus and methods, share a common objective of elucidating the human experience, albeit through different lenses. The literature reviewed in this section delves into the traditional methodologies employed in these fields, examining their foundational principles, applications, and the inherent limitations that have prompted the emergence of more holistic approaches.

The evolution of personality assessment techniques has transitioned from ancient methods like humorism, phrenology, and physiognomy, which correlated physical attributes with personality traits, to more empirical and scientifically grounded approaches (Cheung et al., 1998; Craik, 1986). A significant development in this field was Sir Francis Galton's exploration of the lexical hypothesis in the late nineteenth century, paving the way for systematic and language-based analysis of personality traits (Butcher, 2010). The twentieth century marked a further shift towards dimensional measures of personality, as seen in the introduction of projective techniques like the Rorschach test and the development of self-report inventories, epitomized by the Minnesota Multiphasic Personality Inventory (MMPI) (Butcher, 2010). In recent times, personality assessment has diversified to include a variety of empirically grounded methods, such as self-report inventories, projective tests, behavioral observation, and brain imaging techniques, collectively aiming to provide a comprehensive understanding of individual personality traits, with an emphasis on the correlation of brain activity with these traits (Mosca et al., 2023; Selzer et al., 1987).

Self-report inventories, a key component of psychometric tools in personality assessment, require individuals to respond to statements about themselves, aiding in the assessment of various personality traits. This widely used method, based on the assumption that individuals can accurately self-describe, has been extensively studied for its validity and utility across different contexts and populations (Boyle et al., 2014; Christensen et al., 2019). Complementing these are projective techniques, such as the Rorschach Inkblot Test and the Thematic Apperception Test, which use ambiguous stimuli to elicit responses that reveal deeper layers of personality, including unconscious motives and conflicts (Wiggins et al., 2003; Bender et al., 2014). Behavioral observation, another approach, assesses personality traits by observing and recording behavior in various situations, particularly useful where direct self-reporting is not feasible (Schröder et al., 2021; Roberts et al., 2011). The field also includes observational measures, ratings by others, and performance tests, emphasizing the need for a holistic view of personality through a combination of methods (Al-Dajani et al., 2016; Goodstein et al., 1999). In organizational psychology, psychometric tools like cognitive ability tests and situational judgment tests are crucial in hiring processes, offering objective insights into candidate suitability and predicting job performance (Black, 2000; Mussel et al., 2016). Overall, the advancement and refinement of these tools, from self-report inventories to behavioral observations, have provided nuanced insights into personalities, demonstrating their importance in both clinical settings and areas like recruitment and talent management.
The advent of digital technology has significantly contributed to personality profiling through the emergence of digital footprints. Studies by Lambiotte (2014) and Buraya (2018) underscore the potential of using digital records, such as social media activity and online interactions, for inferring personality traits. These digital footprints provide insights into private traits, motivations, and emotional makeup, enhancing the accuracy of personality profiling by leveraging the continuous data stream generated in digital spaces. However, caution is advised in interpreting these records. Alison (2002) highlights the need for a nuanced understanding of the relationship between behavior and background characteristics and warns against overreliance on simplistic trait approaches, especially in contexts like criminal investigations, due to ethical and legal implications. In addition to digital records, accomplishment records like work performance evaluations are also valuable for personality assessment, aiding in predicting job performance and satisfaction (Harris, 1982; Miller et al., 2015). These findings indicate that both digital and accomplishment-based records are instrumental in understanding and predicting an individual's personality traits in various contexts, with a strong emphasis on ethical considerations in their usage.

Classical methods in historical research, essential for understanding and interpreting the past, have evolved significantly to adapt to the diversification of scholarship and growth in usable sources. These methods, including palaeography, diplomatics, chronology, and epigraphy, are foundational in historical sciences, aiding in the evaluation and analysis of source materials, as elucidated by Grigg (1991) and further developed by Friedrich et al. (2017). The philological method, pivotal in classical studies, has been central since the late nineteenth century, enabling a deeper understanding of historical texts through the interpretation of ancient texts and inscriptions (Kazarov, 2019). Wadhwani (2017) highlights the application of these methods in organization and management studies, using abductive reasoning to analyze historical data from a contemporary perspective. Historians like Wolf (2018) have expanded their research approach to include a wide array of primary and secondary sources, including oral histories, ensuring narratives are constructed based on accurate and reliable historical evidence. These classical methods, evolving alongside modern research approaches, underscore the dynamic nature of historical research and its critical role in accurately recapturing the complexities of the past.

Biographical writing in history utilizes various techniques and approaches to enrich narratives of individual lives. It typically employs a narrative structure that chronicles the subject's life from birth to death, interweaving historical context and, where necessary, posthumous events. The research process is comprehensive, gathering data from diverse sources like anecdotes, interviews, letters, public records, and historical documents, often with authors documenting their process to provide insights into their techniques and the challenges faced (Krasny, 2006; Renders, 2013). Some biographies are presented as historical fiction, blending factual information with fictional elements for more engaging narratives. Collective biographies offer broader perspectives on historical events and societal dynamics, presenting a comprehensive view of the societal and historical contexts of individuals or groups (Rollyson, 2013; Evans & Reynolds, 2012). The biographical approach uses life stories to understand broader historical contexts and societal changes, illustrating the impact of individual lives on historical events (Meister, 2018; Ambrosius, 2004). Incorporating narrative elements like character development, setting, and plot progression, biographical writing provides a deeper understanding of subjects' lives and historical contexts, offering valuable insights into the lives of historical figures and influential individuals (Mann,
Biographical writing thus plays a vital role in understanding and interpreting the past, creating a direct connection to historical periods and individuals.

Narrative research in historical biography, grounded in existential phenomenology, symbolic interactionism, feminist, and psychoanalytic perspectives, uses storytelling to interpret human experiences within historical contexts (Fabiś et al., 2017; Mitchell & Egudo, 2003). This approach, providing insights for both academic and non-academic audiences, balances the focus on individual lives with broader historical questions, allowing for a nuanced understanding of the interplay between personal narratives and larger historical forces (Meister, 2018; Hibbard, 2006). Historical fiction and collective biographies offer unique perspectives in biographical writing; historical fiction blends factual information with fictional elements for compelling narratives, while collective biographies provide broader perspectives on historical events and social dynamics, exploring the complexities of historical periods and societal changes through individual and collective experiences (De Temmerman & Demoen, 2016; Mann, 2009). Narrative structure and techniques, crucial for crafting engaging biographies, typically involve chronological order with character development, setting, and plot progression, making stories engaging and conveying subjects' life journeys coherently. This approach contrasts with traditional narrative history's focus on chronological order and events, while modern narrative history emphasizes structures, general trends, and regional variations (Stone, 1979; Oldfield, 2012).

Research on biographical methods has thoroughly explored their effectiveness across various contexts, demonstrating diverse applications and insights. Studies by Aamodt (1987) and Telonson (1983) have compared different methods for scoring biographical information, with Aamodt finding the vertical percent method to be effective due to significant validity coefficients, while Telonson highlighted the superiority of the rare response method. Chamberlayne (2000) and Smith (2012) have offered comprehensive perspectives on biographical research, emphasizing its historical and philosophical origins and its importance in pedagogical methodology, respectively. This research underlines the utility of biographical methods in shedding light on societal transformation processes and understanding the influence of society on individual development in pedagogical sciences. These methods have found successful applications in fields like applied linguistics and migration studies, providing insights into life trajectories, individual experiences, and interactive subjectivity in the context of migration and societal integration (Iosifides, 2009; Artamoshkina et al., 2021). Overall, the exploration of biographical and historical research methodologies reveals their critical role in revealing the complexities of individual lives within broader historical contexts, bridging the gap between personal narratives and macro-level historical processes. This review highlights the dynamic interplay of narrative structures, the integration of diverse sources in biographical writing, and the effectiveness of using digital footprints in personality profiling and critical narrative approaches in historical biographies.

However, these methodologies are not without limitations. The reliance on digital records in personality profiling raises significant ethical and privacy concerns, necessitating careful consideration and balance. In historical biography, the blending of fact and fiction, while enhancing narrative appeal, can blur the lines between historical accuracy and creative interpretation. The critical narrative approach, aiming to strike a balance between individual stories and broader historical themes, sometimes struggles to maintain objectivity and can be subject to the biases of the author.

http://ipublishing.intimal.edu.my/jods.html
Results and Discussion

In light of the comprehensive review of biographical and historical research methodologies, it becomes evident that while traditional approaches have provided foundational insights, there is a burgeoning need for more nuanced and multi-dimensional methods. The preceding analysis underscores the necessity to transition from conventional techniques, often constrained by limited datasets and an overemphasis on notable events, to a more holistic approach that captures the full spectrum of human experiences. This section thus aims to address the identified gaps and limitations, proposing strategic enhancements to biographical profiling and historical research methodologies. Emphasizing the integration of diverse archival materials, leveraging advanced technology, and maintaining ethical rigor, these recommendations are designed to foster a more intricate and authentic representation of individuals within their historical contexts. The goal is to establish methodologies that not only recount the grandiose events but also illuminate the mundane yet significant aspects of individual lives, thereby enriching our understanding of both the past and the individuals who shaped it with The Holistic Archival Personality Profiling Model (HAPPM) (Table 1).

Table 1. The Holistic Archival Personality Profiling Model (HAPPM)

<table>
<thead>
<tr>
<th>Component</th>
<th>Description</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data Source Diversification</strong></td>
<td>Inclusion of a wide range of archival materials like personal correspondences, social media footprints, and family memorabilia.</td>
<td>To capture a more nuanced view of individuals, encompassing everyday experiences and personal interactions.</td>
</tr>
<tr>
<td><strong>Digitization and Data Transformation</strong></td>
<td>Conversion of various data forms, including handwritten documents, into machine-readable formats.</td>
<td>Facilitates the extraction and analysis of comprehensive data, preserving historical documents and enhancing accessibility.</td>
</tr>
<tr>
<td><strong>Semantic Data Classification</strong></td>
<td>Use of biotags, chronotags, and geotags for organizing data within specific spatial and temporal contexts.</td>
<td>Enhances data organization and contextualization, providing a detailed understanding of subjects’ life experiences.</td>
</tr>
<tr>
<td><strong>Advanced Language Models</strong></td>
<td>Employment of large language models for analyzing digitized and tagged data.</td>
<td>To simulate conversations with historical figures, offering a holistic representation of personalities.</td>
</tr>
<tr>
<td><strong>Innovative Data Analysis Techniques</strong></td>
<td>Development and application of advanced data analysis techniques for interpreting extensive data.</td>
<td>Ensures a more accurate representation of individuals’ lives by uncovering patterns and insights from complex datasets.</td>
</tr>
<tr>
<td><strong>Ethical and Privacy Considerations</strong></td>
<td>Establishment of guidelines for ethical use of personal data and maintenance of transparency in data handling.</td>
<td>Respects privacy and consent, ensuring ethical research practices.</td>
</tr>
<tr>
<td><strong>Interdisciplinary Collaboration</strong></td>
<td>Collaboration between historians, data scientists, and technologists.</td>
<td>Enriches biographical research with diverse perspectives and expertise.</td>
</tr>
<tr>
<td><strong>Training and Capacity Building</strong></td>
<td>Provision of training in digital tools and data analysis methods.</td>
<td>Enhances researchers’ skills in modern research methodologies.</td>
</tr>
<tr>
<td><strong>Continuous Methodological Evaluation</strong></td>
<td>Regular evaluation and refinement of methodologies through feedback and iterative improvement.</td>
<td>Ensures methodologies remain relevant, effective, and ethically sound</td>
</tr>
</tbody>
</table>
The research has yielded insightful results, demonstrating the effectiveness of the employed methodologies, and has offered a novel perspective on biographical and historical narratives. The inclusion of a broad range of archival materials, such as personal correspondences, social media footprints, and family memorabilia, has significantly enriched the research canvas. These diverse data sources have unveiled nuanced perspectives of individuals, capturing their everyday experiences and interactions which are often overlooked in traditional datasets. This holistic collection of data sources necessitated a reevaluation of standard research practices, prompting scholars to extend beyond conventional archives into personal and digital realms.

The digitization and transformation of various data forms, including handwritten documents, into machine-readable text, has proven to be a pivotal step. This process has preserved the longevity of fragile historical documents and has opened new avenues for comprehensive data analysis. It has enabled the application of advanced analytical tools and technologies, facilitating the extraction and interpretation of complex patterns within large datasets. Also, the strategic use of biographical tags - biotags, chronotags, and geotags - has introduced a refined method of semantically classifying data within specific spatial and temporal contexts. This system of tagging has effectively transformed raw data into a structured format, simplifying the complex task of historical analysis. It has enhanced the organization and contextualization of information, providing a more detailed understanding of the subjects' life experiences.

The employment of LLM has been a significant technological advancement in analyzing digitized and semantically tagged data. These models have enabled researchers to simulate conversations with historical figures, offering a more holistic representation of their personalities. Additionally, modern data analysis techniques, including pattern recognition, thematic analysis, and sentiment analysis, have been crucial in interpreting the extensive data gathered from diverse sources. At the same time, throughout the research process, careful attention has been paid to ethical and privacy concerns, especially when dealing with personal data from digital sources. The establishment of clear guidelines for the ethical use of personal data has been paramount. This has involved obtaining permission for the use of personal information and being sensitive to the context in which the data was originally shared. Transparency in data handling has also been a key focus, ensuring the integrity of the research.

The study has embraced interdisciplinary collaboration and enhanced skill sets through training and capacity building. This collaborative environment has led to innovative approaches in data analysis, richer interpretations of historical narratives, and more nuanced representations of personalities. The development of these skills has enabled researchers to more effectively harness the power of digital technologies in their work. A commitment to continuous methodological evaluation has been integral to this research. Regular and critical assessments of current research methodologies have been conducted, focusing on both the efficacy and ethical implications of the methods. The establishment of robust feedback mechanisms has facilitated iterative improvements in research practices. Therefore, the integration of a multi-disciplinary approach in biographical and historical research, as demonstrated in this study, represents a significant advancement in the field. The methodologies employed have not only facilitated a more intricate and authentic representation of individuals but have also redefined the boundaries of historical and biographical research. This research underscores the importance of expanding traditional methodologies to

http://ipublishing.intimal.edu.my/jods.html
more inclusive and detailed explorations, capturing the full spectrum of human experiences and offering a comprehensive picture of historical narratives.

**Conclusion**

The in-depth exploration conducted in this article underscores the critical need for a paradigm shift in the realm of biographical profiling and historical research. Traditional methodologies, while laying a foundational framework, have predominantly been constrained by their dependence on fragmented and often superficial datasets, primarily focusing on official records and notable events. This conventional approach has frequently resulted in a limited portrayal of individuals, overlooking the rich diversity of personal experiences and interactions that profoundly shape human personalities. In response to this, the Holistic Archival Personality Profiling Model (HAPPM) was proposed, marking a significant advancement with its multi-disciplinary approach that encompasses an extensive array of archival materials, cutting-edge technology, and innovative analytical techniques.

The recommendations for implementing HAPPM highlight the importance of diversifying data sources, incorporating not only personal correspondences and family memorabilia but also digital footprints like social media activities. The model emphasizes the digitization and semantic classification of data, utilizing biotags, chronotags, geotags, and potentially new tags like sociotags and polititags. Sociotags could provide insights into socio-economic backgrounds, delineating distinctions such as wealthy/poor or privileged/oppressed, while polititags might offer perspectives on political affiliations or leanings. These additions would further enrich the model, allowing for a more comprehensive and nuanced understanding of individuals in their historical and social contexts. The application of LLMs and sophisticated data analysis techniques within HAPPM promises a more profound insight into personalities, simulating conversations with historical figures and providing a more holistic representation. Additionally, the model is cognizant of ethical considerations, advocating for the responsible and transparent use of personal data.

Looking ahead, the journey of refining and enhancing HAPPM is ongoing. Future research should focus on applying this model in varied historical and biographical contexts, assessing its effectiveness, and ensuring its adaptability to encompass additional tags like sociotags and polititags. This process will involve continuous methodological evaluation, interdisciplinary collaboration, and dedicated training for researchers. The approach signifies a pivotal step towards redefining historical and biographical research. By embracing a more inclusive, detailed, and ethically conscious approach, this model aims to capture not only the grandiose events but also the subtle, everyday aspects of individual lives. It holds the promise of painting a fuller, more authentic picture of personalities, significantly enriching our understanding of the intricate web of human experiences through time and space.

**Data Availability**
Data available upon request.
Conflicts of Interest
The authors declare that there is no conflict of interest regarding the publication of this paper.

Funding Statement
NA

Authors’ Contributions
Conceptualization, P. Ellsworth; Methodology, P. Ellsworth; Validation, J. Hutson; Investigation, J. Hutson – Original Draft Preparation, J. Hutson; Writing – Review & Editing, J. Hutson.; Visualization, J. Hutson.

References


Suthar, A. B., Khalifa, A., Yin, S., Wenz, K., Ma Fat, D., Mills, S. L., ... & Mrkic, S. (2019). Evaluation of approaches to strengthen civil registration and vital statistics systems: A systematic review and synthesis of policies in 25 countries. PLoS Medicine, 16(9), e1002929.


