MIS581 Capstone Project: Literature Review

Nicholas Laeder

Colorado State University - Global Campus

MIS581 - Data Analytics Capstone

Steve Chung

6/25//2025

Email marketing remains an indispensable channel for small to medium-sized organizations (SMEs) to engage customers, promote services, and drive conversions. These platforms, such as Constant Contact and Mailchimp, generate a vast amount of data on subscriber interactions and campaign performance. However, a significant problem is the underutilization of this email marketing data by SMEs, primarily due to a lack of resources and technical expertise required to extract actionable insights. This often results in missed opportunities for campaign optimization, improved targeting, and increased marketing return on investment.

This capstone project seeks to address this critical gap by designing and developing a lightweight, cloud-based data warehousing proof of concept to integrate email marketing data from Constant Contact and Mailchimp. A novel aspect of this solution is the demonstration of Large Language Models (LLMs) to analyze data and generate actionable insights in an accessible, conversational "podcast" format, empowering non-technical business owners to make better-informed decisions . The central research question guiding this endeavor is: "For small businesses utilizing email marketing platforms (such as Constant Contact and Mailchimp), what are the key characteristics of high-performing email campaigns, and how do identifiable subscriber engagement patterns influence desired business outcomes (e.g., conversions, retention, new sign-ups)?". This literature review will delve into existing scholarly work on data warehousing, business intelligence (BI) adoption challenges in SMEs, and the evolving role of AI in analytics, providing a theoretical foundation and practical context for the proposed capstone solution.

Challenges in SME Data Utilization and Business Intelligence Adoption

SMEs consistently struggle with data processing due to insufficient infrastructure and technical expertise (Dötlinger et al., n.d.). Scholarly work indicates that SMEs are late adopters of Business Intelligence (BI) solutions, lagging behind larger enterprises in leveraging BI's potential (Gudfinnsson, 2019). Gudfinnsson (2019) identifies several key challenges for BI adoption in SMEs, including a general lack of knowledge on BI capabilities, limited interest from management, a tendency to rely on "gut-feeling" over data, and difficulties with outsourced IT support (Gudfinnsson, 2019). SMEs frequently lack the financial resources, specialized skills, or extensive experience needed to implement complex BI systems (Gudfinnsson, 2019).

Beyond general BI, the adoption of Self-Service Business Intelligence (SSBI)—which aims to empower non-technical users directly—introduces its own set of user-related challenges (Lennerholt et al., 2021). Lennerholt, Van Laere, and Söderström (2021) highlight that while SSBI is intended to remove bottlenecks by reducing reliance on power users, new dependencies emerge (Lennerholt et al., 2021). Casual users face difficulties in accessing, locating, and properly using data sources, often still requiring IT support to add missing data (Lennerholt et al., 2021). Furthermore, users exhibit limited competence levels and general IT skills, finding SSBI tools difficult to use, which leads them to create isolated and potentially faulty analytical solutions (Lennerholt et al., 2021). Critically, there are significant challenges in assuring the quality of SSBI reports and a notable lack of governance for these user-generated reports, leading to redundant or inaccurate information that can negatively impact decision-making (Lennerholt et al., 2021). The authors emphasize the need for formal education that goes beyond mere tool proficiency, focusing instead on "moving toward a new way of working" to integrate SSBI skills into daily routines (Lennerholt et al., 2021). These user-centric challenges underscore the necessity of the capstone project's focus on delivering insights in an accessible and easily consumable format to its non-technical SME audience.

Cloud Data Warehousing as a Strategic Solution for SMEs

Cloud computing has revolutionized data management by eliminating the burden of building and maintaining costly hardware and software infrastructure (Yasser & Alserafi, 2023). Its pay-as-you-go model and vast scalability make it particularly appealing for SMEs seeking a systematic approach to data analysis without significant upfront capital investment (Dötlinger et al., n.d.; Yasser & Alserafi, 2023). Yasser and Alserafi (2023) propose key criteria for evaluating cloud-based Data Warehouses (CDWs), including ETL processes, data querying, data visualization, and crucial security aspects like encryption, access control, audit logging, backups, and disaster recovery (Yasser & Alserafi, 2023). Their research specifically explores Google BigQuery as a cost-effective, serverless data warehouse solution that incorporates machine learning and BI tools, capable of scaling dynamically with user data (Yasser & Alserafi, 2023). The capstone project's architectural plan, which utilizes Google BigQuery as its central data warehouse for transformation and analytical querying, aligns directly with these established benefits and evaluation criteria.

The concept of a "lightweight" data warehouse for SMEs is further supported by Dötlinger, Penz, Reiter, and Widauer (n.d.), who note that cloud-based DWH services are typically designed as Software as a Service (SaaS) with flexible, consumption-based pricing models that directly target the needs of small and medium businesses (Dötlinger et al., n.d.). While comparing various services, they highlight the importance of ease of setup and suitability for specific SME requirements (Dötlinger et al., n.d.). The choice of Google BigQuery in the capstone project is further reinforced by a comparative study by Salqvist (2023), a master's thesis that compared Data Warehouses (specifically BigQuery) against Data Lakehouses (like Delta Lake) (Salqvist, 2023). Salqvist found that BigQuery consistently demonstrated significantly lower average latency (i.e., faster performance) than Delta Lake across various data scales when executing Decision Support System (DSS) queries (Salqvist, 2023). This finding validates BigQuery's strong performance capabilities for analytical workloads, making it a robust choice for the capstone's objectives.

Augmented Analytics and AI for Accessible Insights

The integration of artificial intelligence (AI) is profoundly transforming the landscape of data management and analytics (Mally, 2023). Mally (2023) explains that AI analytics, utilizing sophisticated algorithms and machine learning, enables the interpretation, forecasting, and application of data to shape future strategic planning and decision-making (Mally, 2023). This synergy between AI and cloud data warehousing amplifies the capabilities of cloud platforms, introducing predictive and prescriptive analytics and transforming data into a dynamic instrument for strategic development and competitive advantage (Mally, 2023).

While AI is capable of generating basic narrative reports, it often leads to oversimplification if the narrative element is neglected, creating visually appealing but weak data stories (Dykes, 2024). Forbes (Dykes, 2024) posits that the future of data storytelling is "augmented, not automated," emphasizing a collaborative approach where human insight complements AI's analytical power to deliver more compelling and impactful data stories (Dykes, 2024). This perspective perfectly aligns with the capstone project's innovative approach of using LLMs to generate "podcast-style" narratives, aiming to provide comprehensive and accessible insights to non-technical business owners (Laeder, 2025a; Laeder, 2025b). The practicality of this delivery method is reinforced by tools like Jellypod, which demonstrate the feasibility of converting existing text content into natural-sounding speech using AI voices, simplifying the production and promotion of audio content (Jellypod, n.d.).

Conclusion

The literature overwhelmingly supports the critical need for solutions that enable SMEs to better utilize their data for strategic decision-making (Dötlinger et al., n.d.; Gudfinnsson, 2019). The challenges faced by these businesses—ranging from a lack of technical expertise and infrastructure to difficulties with self-service tools and ensuring data quality—underscore the value of a lightweight, cloud-based data warehousing solution (Dötlinger et al., n.d.; Gudfinnsson, 2019; Laeder, 2025b; Lennerholt et al., 2021). The capstone project's choice of Google BigQuery is well-supported by its demonstrated scalability, cost-effectiveness, and strong analytical performance (Salqvist, 2023; Yasser & Alserafi, 2023). Furthermore, the innovative integration of AI (LLMs) to transform complex data into easily digestible, podcast-style insights directly addresses the user-related challenges of traditional and self-service BI, bridging the gap for non-technical business owners (Dykes, 2024;). By building on the foundations laid by this scholarly literature, the project aims to offer a practical, robust, and accessible proof of concept, contributing to both the academic understanding and practical application of data analytics in the SME landscape.

References

Dötlinger, L., Penz, M., Reiter, M., & Widauer, S. (n.d.). *Lightweight DWH Data Analysis for SMEs*. University of Innsbruck, Austria.

Dykes, B. (2024, February 27). *The Future Of Data Storytelling Is Augmented, Not Automated*. Forbes.

Gudfinnsson, K. (2019). *Towards facilitating BI adoption in small and medium sized manufacturing companies* (Doctoral dissertation, University of Skövde).

Jellypod. (n.d.). *Grow Your Business with AI-Generated Podcasts*.

Lennerholt, C., Van Laere, J., & Söderström, E. (2021). User-Related Challenges of Self-Service Business Intelligence. *Information Systems Management, 38*(4), 309–323.

Mally, P. K. (2023). Cloud Data Warehousing and AI Analytics: A Comprehensive Review of Literature. *International Journal of Computer Trends and Technology, 71*(10), 28–38.

Salqvist, P. (2023). *Abstract: This thesis aimed to assess a given Data Warehouse against a well-suited Data Lakehouse in terms of read performance and scalability*. (Master's thesis, KTH Royal Institute of Technology).

Yasser, M. F., & Alserafi, M. M. (2023). Cloud-Based Data Warehousing Solutions Capabilities and Challenges. *2023 International Conference on Computer Science and Information Technology (CSIT)*, 1–6.