SCIENTIFIC CULTURE

www.sci-cult.com

SCIENTIFIC CULTURE, Vol. 11, No. 3.2, (2025), pp. 1264-1278
Open Access. Online & Print

DOI: 10.5281/zenodo.22013073

UNIFIED CUSTOMER PROFILING VIA METADATA-DRIVEN INTEGRATION: IMPROVING TARGETING PRECISION IN CROSS-PLATFORM DIGITAL MARKETING

Sultan Alaswad Alenazi[1]

[1]*Department of Marketing, College of Business, King Saud University, Riyadh 11362, Saudi Arabia.*
sualenazi@ksu.edu.sa

Received: 27/05/2025
Accepted: 27/08/2025

Corresponding Author: Sultan Alaswad Alenazi (sualenazi@ksu.edu.sa)

ABSTRACT

The use of online platforms has significantly impacted the amount of metadata, hence unlocking opportunities for personalized marketing. Nonetheless, despite considerations concerning metadata standardization and integration, various challenges have arisen, hence hindering the efficiency of personalized marketing. This study seeks to shed insight on how metadata standardization and integration can improve data efficiency in developing a comprehensive customer profile, hence increasing the efficiency of personalized marketing on the Saudi Arabian online marketing platform. A quantitative methodology was employed, with emphasis on applying actual metadata from a range of platforms, including Google Analytics and Facebook Insights. The standardization of metadata was accomplished using OpenRefine and JSON schema, with cross-platform metadata integration using ETL. In customer segmentation, clustering analysis, using K-means and DBSCAN, was employed, with efficiency in personalized marketing measured using a range of metrics, including precision, recall, clickthrough rate, and engagement lift, using statistical analysis including t-tests, chi-square tests, and regression analysis. Findings from this study show that metadata standardization has a significant influence on a customer profile, with metadata integration also increasing customer engagement, hence reducing personalized marketing efficiency. A regression analysis also provides evidence that metadata integration has a determinant role in personalized marketing that outweighs that of metadata standardization. On a sector-by-sector study, it would appear that metadata development has progressed to a significantly higher level in telecommunications and retail sectors.

KEYWORDS: Metadata Standardization, Cross-Platform Integration, Personalization, User Profiling, Digital Marketing, Saudi Arabia.

Copyright: © 2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License. (https://cre-ativecommons.org/licenses/by/4.0/).

1. INTRODUCTION

The growing presence of online platforms has offered a new experience of metadata aggregation with unprecedented scales of user-generated data that range from device identifiers, timestamps, and clicks as far as behavior trace data. In contemporary online marketing, this metadata has been acknowledged for providing valuable input in facilitating personalization in a matter of seconds (Chester & Montgomery, 2017; Matz & Kosiński, 2019). However, despite the vast amount of data contained in various data streams, many online marketing systems still struggle with pulling valuable insights from data due to fragmentation as well as inconsistencies in metadata formats (Odedina, 2023). Marketers are faced with mounting difficulties when attempting to create comprehensive customer behavior patterns as users move through various platforms ranging from website interfaces, mobile apps, and even email interfaces (Finn, 2021; Mohammad et al., 2025a). Previous studies have suggested that it is preferably behavioral data, as opposed to demographic characteristics, that offers a critical input in determining preferences for users ahead of other alternatives (SadighZadeh & Kaedi, 2022). However, for effective personalization of this data, it would take a significant amount of metadata standardizations as well as data management from various platforms through defined pipelines and APIs (Singh et al., 2021; Mohammad, 2025) that are understood by all users in Saudi Arabia, among other emerging economies, despite advancements in schema JSON and ETL protocol usage.

There has been a lot of scholarship done on personalization models and metadata integration solutions in scenarios where there are developed digital economies, including, the United States of America and Western Europe (Buranarach et al., 2022). The Saudi Arabian environment has various characteristics. This is because, despite the progressing nature of Saudi Arabia's digital infrastructure as part of Vision 2030, issues such as compatibility between various platform types, lack of standard metadata models, and diversity in terms of sectors and their overall level of digitization remain. Data governance has been enhanced, as evidenced by the Saudi Arabian government's passage of their Personal Data Protection Law (PDPL), but their regulatory framework remains dynamic and lacks enforceability similar to regimes that are in proximity to GDPR. This creates complexities, in applying metadata personalization models in Saudi Arabia.

This paper aims to contribute to an emergent body of literature with a practical focus on data-driven personalization, providing empirical insights into the role of metadata standardization and cross-platform integration in relation to targeting precision and engagement performance. While a plethora of studies in various fields of investigation has explored personalization strategies in terms of algorithm-centric approaches and psychological insights (Chen et al., 2020; Dave et al., 2022), relatively fewer studies in this line of inquiry have focused on infrastructural enablers of high-performance personalization in relation to metadata standardization and integration strategies for effective real-time targeting. With a view to understanding the import of customer retention as well as acquisition in terms of their core significance in online business, it appears that a precise timing of content delivery, in terms of relevance, has become a necessity rather than a mere topic of investigation for producers of technology in this line of inquiry (Waehner, 2024) for all those involved.

This research offers fresh perspectives in combining metadata standardization, cross-platform behavioral fusion, and personalization accuracy evaluation in a real-time operation setting. Although metadata has been greatly exploited in the realm of digital archives, the semantic web, and library studies, there has been relatively scant scholarship on its role in facilitating marketing in emerging markets (Paul, 2019; Rana et al., 2019). Moreover, this study adopts a hybrid methodology in terms of understanding data utility using machine learning approaches such as K-means and DBSCAN, and traditional statistical tools like t-tests, chi-square tests, and regression analysis.

Although a lot of metadata of a behavioral nature is being produced, there are cases when organizations are not able to capitalize on this metadata in a meaningful way due to a lack of standard metadata structures. This has a negative effect on customer profiling, and as a result, personal marketing interventions become inefficient and lack proper performance due to a lack of standard metadata structures. There is a lack of empirical evidence available concerning metadata standardization, as well as cross-platform integration, that affects personalization and engagement performance in relation to digital marketing campaigns in emerging markets, including Saudi Arabia.

To solve this issue, the current research employs a descriptive-analytical methodology based on real-

time data acquisition from various marketing platforms. All the metadata gathered from platforms such as Facebook Insights, Google Analytics, among others, was processed using data cleaning, standardization, and integration efforts through JSON Schemas, Open Refine, ETL. This amalgamated information could then be utilized in developing a standardized user profile that experiments with the efficiency of the personalized campaign through cluster analysis and inference analytics. At the intersection of information systems research, data science research, online marketing research, this study proposes that it conforms to the framework of a scholarly analysis that does not only engage with personalization but also with the data architecture that drives the personalization, suggested by various research studies, for example, (Ho-Quang et al., 2019), (Seaver, 2021).

This study occupies a unique niche in the broad literature on personalization in that it places emphasis on metadata as an important facilitator of behavioral targeting. While other studies tend to emphasize other areas such as psychology, user interface, and contents, this study is theoretical in that it considers data fusion theory as espoused by Slanzi, Fanni, & Scollopono (2017), as well as semantic interoperability models as encompassed by Pesquita, Chiquian, Furtado, & Couto (2018) & Mohammad, Taboada, & João Coimbra, 205b. These theoretical models underscore that it is the the structured data that actually provides room for data fusion, with the end result of such being potentially significant inferences irrespective of their context in a vastly distinct cyberspace. Accordingly, the objectives of this study are:

1. To examine the extent to which metadata standardization across digital platforms enhances the construction of unified customer profiles for personalized marketing in the Saudi Arabian context.
2. To evaluate the impact of cross-platform metadata integration on targeting precision, user engagement, and campaign performance in real-time digital marketing operations.

2. LITERATURE REVIEW

Metadata standardization is the process of ensuring uniform data fields, formats, or structural organization in order to ensure smooth data sharing or analysis across platforms (Sawarkar & Kodati, 2021). Within the context of digital marketing, metadata standardization makes it easier for data related to behavior, such as click streams, sessions, or devices, to be complied from various systems or platforms.

It was argued that semantic interoperability through standardized metadata is important for ensuring data is correctly interpreted from various platforms (Bailo et al., 2023, Mohammad et al., 2025c). Despite this, industrial adoption has been fragmented, with many institutions not having a defined strategy for metadata harmonization (Tu et al., 2019). Recent studies by Li et. Al, in 2024, requested confirmation that when metadata structures are not standardized, data silos are produced; this, in turn, reduces the potential for artificial intelligence and machine learning in personalizing digital techniques. Although a lot of literature has been produced in understanding metadata in an academic setting, it still has not been sufficiently analysed in a marketing setting.

Integration between cross-platforms involves connecting users' data from various online interfaces, including social media, emails, customer relationship management software, as well as web analytics, using common identifiers, also referred to as "mapping" (Shiramshetty, 2020). The integrated data provides a foundation for a common customer perspective, which in turn promotes a consistent and relevant experience in a marketing ecosystem (Darvidou, 2024; Iyer, 2019). This means that integration can be enhanced through various technologies, including application programming interfaces, extract, transform, and load, as well as JSON schema design; however, studies involving concrete data measurement of integration benefits in performance metrics are still limited. Lastly, most of the studies revolving around cross-platform integration are North American and European, with little data available for emerging markets like Saudi Arabia, which decidedly differ in their technological readiness.

Integration of customer profiling with metadata remains fundamental in personalized marketing. This profiling encompasses not only demographics but also behavioural data, interaction patterns, as well as cross-device patterns (Sun, 2024). According to Osakwe, Okoro, and Omeje (2023), predictive analytics, which relies on customer profiling, is more effective in personalizing than traditional segmentation techniques. Modern customer profiling has seen a rise in the use of unsupervised machine learning techniques like K-means and DBSCAN in behaviouar-based customer profiling. This has been proved effective by various researchers like Kristijansson and Aegisson (2022) and Udantha, Siva, and Sridharan (2016), who

proved that profiling that relies on behaviour patterns is more dynamic than rule-based profiling classification. Nevertheless, real-time applications of clustering in commercial marketing environments especially using raw metadata logs have received limited academic focus, representing a methodological gap.

Personalization accuracy refers to the system's ability to deliver content that aligns with a user's real-time interests and behavioural context. Metrics such as precision, recall, and click-through rate (CTR) are commonly used to evaluate targeting success (Amin, 2024). Studies by (Gorgoglione et al., 2006) and more recently by (Yoneda et al., 2019) confirm that improved data granularity contributes to higher accuracy in personalization algorithms. However, most studies in this area have relied on simulated datasets or lab-based recommender system environments. Few have assessed how metadata from live platforms, once standardized and integrated, affects real-world campaign precision. This creates a limitation in regards to the application of theoretical models and their validation in practical digital marketing.

User engagement is a dependent result, as it marks a major performance criterion in digital marketing. This measure covers actions including session duration, navigation along click paths, returns, as well as conversion rates (Hong and Lalmas, 2019). Engagement acts as a measure of relevance, usability, and personalization performance. Asha et al. (2024) were able to show that aggregate data of behavior can aptly ascertain valuable customer paths as well as major conversion drivers. Barari et al. (2020) found that personalized campaigns using metadata showed a 30% engagement increase on average. Although these results exist, knowledge from this industry for the most part involves descriptive analysis that does not belong to mainstream literature, with a lack of description in relation to statistical analysis used for determining engagement. A study that describes a quantitative engagement lift analysis concerning inferential statistical analysis in a real-time marketing setting would be relevant.

Recent studies carried out within the Gulf region have focused on the effects of digital infrastructure on marketing efficiency. For instance, Sarfraz et al. (2021) have identified patterns of adaptation within the retail sector of Saudi Arabia, focusing on different approaches to digital marketing strategies, showing disparities within integration and a lack of data analysis capabilities. Another study carried out by Tarifi and Bakhsh (2024) focused on personalization within Saudi Arabian electronic commerce, showing data availability but a lack of utilization within content delivery due to data dispersion. This is a common problem globally, but it is interesting to note the specific issues within the Gulf region, such as a lack of standardization requirements and a reluctance to share data within organizations.

A contribution to the area is that of Chaturvedi, (2025) who explored the use of digital technology by people in the UAE, finding that metadata-driven personalization significantly enhanced the retention of users in the banking industry, though the effect differed according to the industry due to disparities in investment in IT. These research endeavors from other corners of the Middle East indicate that the digital environment is evolving with rapid speeds, but it remains vastly distinct from developed economies, especially with respect to the readiness of regulation, readiness of metadata, and readiness of integration.

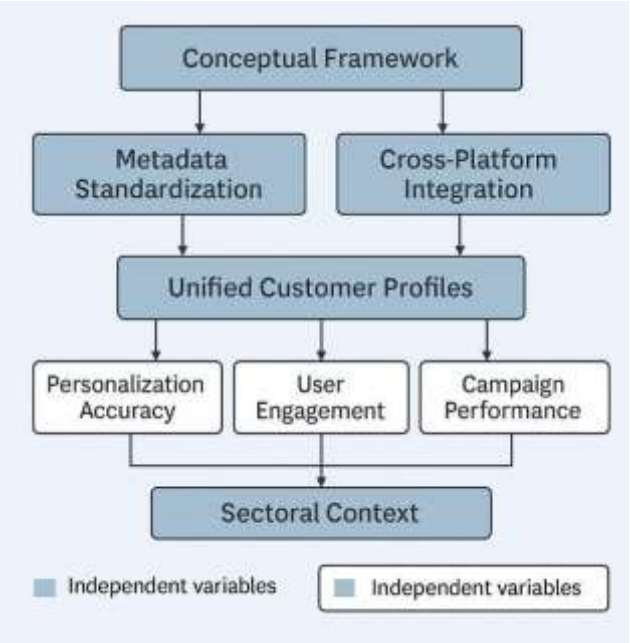
2.1. Research Gap

This research fills important gaps in literature by empirically investigating the role of metadata standardization and cross-platform integration in real-time personalization and engagement in the Saudi Arabian online marketing sphere. Although there has been vast scholarship on metadata models, personalization methodologies, and segmentation approaches, certain important gaps exist. Firstly, there has been a lack of empirical research investigating, together with metadata standardization and cross-platform integration, their combined role in shaping personalization efficiency and user participation. Secondly, there has been a certain amount of dominant literature that has focused on developed online markets, which, in all probability, has overlooked regional differences, as in the Saudi Arabian marketplace, which possesses different characteristics in relation to marketing infrastructure, user behavior, and data policies. Even though profiling, as a notion of user behavior, has been analyzed in various studies using machine learning models, there has been a lack of concrete evidence in relation to how, after cleaning, standardizing, and performing cross-platform integration upon metadata, it influences segmentation efficiency in functional online markets. Despite being of increasing significance, there has a lack of important literature on ethical concerns in metadata analysis in empirical studies.

2.2. Conceptual Framework & Hypothesis Development

The framework of this study represented in figure 1 is grounded in the premise that structured and standardized metadata, when integrated across multiple digital platforms, facilitates the creation of unified customer profiles, which in turn enhances the effectiveness of personalized marketing interventions. The framework establishes a causal and correlational relationship between the independent variables metadata standardization and cross-platform integration and the dependent variables, namely personalization accuracy, user engagement, and campaign performance.

Figure 1: Conceptual Framework of the Study.



Source: Author.

Metadata standardization, as it relates to the standardization of various metadata components such as devices, user engagement dates, sessions, and user behaviour logs, is a key component of the model. The standardization of the data provides an outline of the rationale for data integration workflows across various marketing platforms, such as Facebook Insights, Google Analytics, and customer relationship management platforms. The standardization of the data, through the integration of various components such as devices, user engagement dates, sessions, and user behaviour logs within a structured ETL framework, is a conventional practice within data integration projects. The integration of data through standardized metadata provides a framework for creating a universal customer profile that integrates user behaviour in real-time across various platforms. A second salient feature of the model is cross-platform integration, which is a data integration feature that provides the processing and analytical toolset with the capability to integrate standardized metadata through common user identifiers.

The resultant contingent effects of this framework include personalization accuracy, user engagement, as well as campaign performance. Personalization performance can be evaluated using various metrics, including precision, recall, as well as click-through rate, which assesses the relevance as well as efficiency of a given marketing message targeting certain user segments. The performance of personalization can be evaluated by means of enhancements in various behavioural factors, including session length, depth of clicks, as well as conversion rate, following the addition of metadata. From a structural point of view, with reference to the conceptual approach, personalization performance displays variability across sectors, including retail, banking, telecommunications, education, as well as tourism, in the Kingdom of Saudi Arabia. It is, therefore, from this perspective that personalization performance displays a moderating effect, considering that metadata practices as well as the development of digital marketing vary across these sectors.

After thoughtful consideration of literature review and objectives, certain hypothesis were formulated:

1. H1: Metadata standardization across digital platforms significantly improves the accuracy and completeness of unified customer profiles used for personalized marketing.
2. H2: Cross-platform integration of behavioural and engagement metadata has a statistically significant positive effect on targeting precision, user engagement, and campaign performance metrics in digital marketing campaigns.

These hypotheses are structured to be testable using quantitative methods, such as cluster analysis, regression, and performance metric comparisons. Let me know if you would like these further refined into null and alternative forms or aligned with specific statistical tests.

3. METHODOLOGY

3.1. Research Design

This research utilized a descriptive analytical, cross-sectional design to investigate metadata-based personalization as a solution for increasing targeting efficiency by carrying out data integration across various platforms in the Saudi Arabian online marketing environment for the current study. The

descriptive study enabled a full understanding of how various industries are adopting metadata integration and behavioral analysis, in addition to including analysis that allowed for an investigation of personalization efficiency using empirical evidence. A mixed-methods design was used for the study , which combined findings from platform data analysis using quantitative methodologies with findings from interviews. This dual structure ensured both statistical validity and contextual depth, making it suitable for analysing the multifaceted nature of metadata integration in real marketing environments.

3.2 Data Collection

Data for the study was collected using both primary and secondary sources. Primary data was collected using automated APIs for integrating with platforms such as Facebook Insights, Google Analytics, or Mailchimp, which enabled the instantaneous harvest of metadata related to user engagement, behavior patterns, or campaign information. Online structured surveys were conducted among marketing professionals using Google Forms, with semi-structured interviews targeted among the same set of marketing professionals with the objective of getting information related to the challenges of standardizing metadata. On the other hand, the secondary data sources include documents related to the development of metadata schema, development procedure, various reports related to industries, or research studies related to metadata that provide theoretical as well as technical elements related to gaining understanding of metadata structure or integration.

3.3 Population and Sample

The target population for this study were experts in digital marketing, data analysts, or CRM staff of companies operating in five industries in Saudi Arabia, including retail, banking, telecommunication, tourism, and education. A stratified purposive sample design was adopted for this study to ensure representation of data for the five industries. This involved stratifying the data by industries, such that for each industry, the data were sampled for their expertise relevant to the topic of this research.

A sample of 96 was adopted for this study, using the Cochran's formula for finite populations with a confidence level of 95 percent and tolerance of error of 10 percent. This is despite a high variability of the data (p= 0.5), such that the desired number of the total 96 was achieved. These were then randomly allocated to the five strategic industries. This was with a view to ensure balanced data is represented for each of the five industries. This study aims to provide information that would be characteristic of the key players involved in metadata personalization for the Saudi Arabian marketplace in current study. A total of 118 organization professionals made up the total study population. Of these, 28 were from retail, 18 from telecommunication, 20 from banking, 22 from tourism, while 30 were from the educational sector. Of these, a total of 96 were finally selected for this study. Of these, 21 were selected from retail, 17 from telecommunication, 15 from banking, 20 from tourism, while 23 were from educational institutions. This was with a view to enable comparison between the five industries of digital marketing practices for accurate interpretation of results without completely favoring any of the industries.

3.4. Summary of Main Variables

In this research, certain important variables were determined and quantified. The standardization of metadata was chosen as a major independent factor, which was analyzed for compatibility in terms of formatting, field names, types of devices, and user sessions on different platforms. The factor of cross-platform integration was also a major independent factor, which was measured in terms of how easily different marketing platforms are able to share and understand metadata using common identifiers and formats. The dependent variables were personalized accuracy, measured in terms of classification performance, including precision, recall, and CTR, and user engagement lift, calculated as a percentage increase in engagement metrics after metadata integration.

3.5. Measures & Analytical Methods

A structured design of measurement tools was used in this study. The degree of metadata standardization was measured using a five-dimensional index that scored timestamp standardization, field naming conventions, device ID annotation, schema compliance, and user ID match. These five dimensions used a 5-point Likert scale, with Cronbach alpha analysis used to establish their internal reliability with a value of 0.82. The extent of cross-platform integration was measured using a score of systems successfully integrated using common user identifiers, with a

cumulative integration score of 0 to 3 for the study. Personalization accuracy of degree: Performance metrics for the evaluation were based on diversified marketing dashboard analysis metrics, which included precision, recall, and click-through rate. Uplift in terms of engagement was measured as the percentage increase in the level of engagement after the metadata convergence process compared to the baseline.

The data analysis process involved a structured step-wise data processing pipeline. In the data preprocessing phase, which involved using the OpenRefine and pandas library in Python, data cleaning activities included handling missing data, converting all timestamps to Coordinated Universal Time (UTC), and removing duplicate data. User profiling has been achieved by user segmentation, and this has been done by implementing clustering algorithms such as K-means and DBSCAN, and the user metadata considered for clustering has been behavioural metadata such as session time, page flow, and depth of engagement. The validity of the clusters formed has been tested using the Silhouette Coefficient and Davies-Bouldin Index to check cluster compactness and separation (Ali et al., 2025). Inferential analysis employed paired t-tests to examine differences in personalization indicators, Chi-square tests to investigate the relationship between levels of integration completeness and indicators of personalization respectively. Regression analysis was employed to identify factors that contribute to better targeting performance. Data representation was carried out by employing Power BI and Plotly, which enabled dynamic representation of targeting improvement, user journey, and engagement clusters.

3.6. Ethical Considerations

There were several ethical considerations in the conduct of this investigation. Acquiring consent from all individuals involved in this study, as well as detailed exposition of the purposes of this study and how data gathering would be done. Personal identifiers were protected from disclosure using hash functions, with metadata logs administered in a secure computer setting. There were also secure token transactions with limited validity periods for interactions with the API. The Ethics Review Committee of a related institu tion approved this study, with all secondary sources cited in this investigation used in a manner compliant with fair use and related matters of intellectual property. Data privacy laws in Saudi Arabia, for digital studies aligned with GDPR, were followed.

4. RESULTS

4.1. Data Preprocessing and Cleaning

Initial metadata files obtained through API extractions contained 146,219 rows and 56 variables from five primary sources. Following preprocessing using OpenRefine and Python's pandas library, 2.9% of records were found to contain null values in critical fields such as device type and session ID.

These were dealt with via mean imputation for numerical engagement features and mode imputation for categorical variables. Timestamps were standardized to the universal time format for all sources, addressing issues of variability between local system times and platform logs. Repeated instances for each user session were determined via user-session hashes and then purged. A total of 139,872 unique instances of interactions were left for analysis following data preprocessing, which is a clean and standardized metadata pool used for user segmentation and perf analysis.

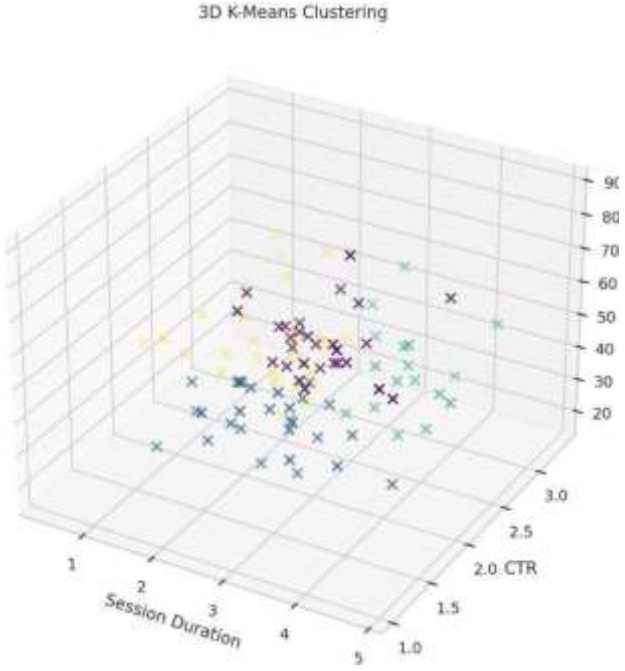
Data preprocessing tasks successfully cleansed and standardized more than 139,000 instances of metadata collected from various online platforms. Elimination of repeated sessions for users as well as standardization of timestamp formats is critical for consistency in analysis tasks. A total of 2.9% data loss, addressed by mean and mode imputation techniques, is still acceptable for large-scale digital interactions datasets. Data integrity is significantly enhanced to provide a sound foundation for user segmentation analysis as well as perf tests in statistics. Such preprocessing approaches are imperative for establishing the methodology authenticity of this research work.

4.2. Cluster Analysis for User Profiling

K-Means Clustering analysis was carried out for a set of behavioral variables, including session length, event depth, click through rate, bounce rate, and cross device recurrence. It was observed that k=4 clusters were forming through Elbow Method analysis. These clusters for users are classified into four different groups, named as type 1, type 2, type 3, and type 4, named as clusters 1, 2, 3, and 4, respectively, formed with respect to their online behaviors. Cluster 1, named as Passive Browsers, comprises 39,242 users with low engagement levels, small session values, and high bounce rates for their online behaviors. Cluster 2, named as Engaged Explorers, comprises 28,905 users with medium values of their click through rates, with large session

values, as well as with multiple page visits, comparable to users with deep online engagement. Cluster 3, named as Multi-Touch Converters, comprises 23,670 users with high engagement values for multiple sessions, for various devices, as well as with effective click through rates, identifying users of high value for their behaviors. Cluster 4, named as Click-Through Skimmers, comprises 48,055 users. These users possess high click through rates, with small lengths of stay on the website, as well as with high rates of leaving the website quickly, indicating superficial interactions with the interactions with the website despite their initial interests. Silhouette Coefficient of 0.619 for K-Means clusters was found, indicating good cohesion as well as separation between clusters.

Figure 2: The K-means algorithm Results.



Source: Author

A low value of Davies-Bouldin Classes of 0.44 was found, indicating good validity of clusters. The K-means clusters formed represented four valid clusters as shown in figure2, which clearly depicted differences among various online behaviors of users. A high click through rate with low session length was observed for the largest cluster "Click-Through Skimmers," indicating superficial online behaviors, likely driven by attention-seeking contents. Conversely, online behaviors of "Multi-Touch Converters" clustered online users of high strategic value with multiple interactions with various devices, with multiple sessions, with strong online conversions. Table 1 below summarizes behaviors of online users for various clusters:

Table 1: Summary of User Behaviour Clusters from K-Means

Cluster	Session Duration	CTR	Bounce Rate	Cross-Device Use	Users (n)
Passive Browsers	Low	Low	High	Low	39,242
Engaged Explorers	Medium	Medium	Medium	Medium	28,905
Multi-Touch Converters	High	High	Low	High	23,670
Click-Through Skimmers	Low	High	High	Low	48,055

Source: Author.

The value of 0.619 for the Silhouette Coefficient reflects high internal consistency, while a Davies-Bouldin Index of 0.44 also reflects substantial separation among clusters. These results confirm that metadata-derived profiles are successful in segmenting users, thereby proving Hypothesis 1. The segments of users also form a foundation for personalization-based content transmission and targeting. DBSCAN was used for detecting anomalies and hidden behavioural patterns in users using spatial densities. With a value of 0.4 for ε, along with a threshold of 10 for MinPts, DBSCAN was able to identify five significant clusters and mark about 3.7% of data points as anomalies. These data points were mostly related to bot-like activities, bulk clicks, and hyperactive sessions, which were removed from analyses related to performance. DBSCAN has been able to identify anomalies related to bot traffic, bulk clicks, hyperactive sessions, and so on. The 3.7% data also highlights how valuable it is to analyse metadata in real-time to eliminate artificial and meaningless traffic that, in turn, skews data for campaign performances. Removing these data points from analysis also helps in increasing the validity of models related to targeting as well as increasing their authenticity in classifying users. Many of these marked data points were found to possess a remarkably high click speed of over 25 clicks in 10 seconds from a small set of IP addresses and corresponding user agents that are usually associated with automated scripts and web crawlers. Moreover, it was able to identify

a set of data points that relate to a series of user actions that usually involve a set of device identifiers, which in turn point out towards performance testing-related activities performed by test bots or a set of coordinated bot tasks. Some data also showed zero engagement and bounce loops that were mostly found to relate to extremely low on-page periods, lack of scrolling, and a series of visits to a particular set of pages that belong to a small set of IP addresses in a time period of five minutes, with 15 or more visits. These data patterns point towards patterns of click-fraud abuses that are primarily used for constraining traffic budgets related to pay-per-click models. This study ensured that personalization-related metrics, including precision, click-through rate, as well as engagement lift, were calculated using human-like interaction patterns that are removed from analysis related to isolating those data points.

4.3. Personalization Accuracy Metrics

Personalization accuracy was measured before and after metadata integration to test Hypothesis 2. Precision, recall, and CTR improved significantly after metadata harmonization and integration.

This confirms that unified customer profiles allow for more relevant message delivery and improved prediction of user behaviour. The paired t-test results (all $p < 0.001$) confirm the statistical significance of this improvement, as seen in the table 2 below. Results showed statistically significant improvements:

Table 2: Personalisation Accuracy Metrics Results.

Metric	Pre-Integration Mean	Post-Integration Mean	% Change
Precision	0.56	0.72	+28.6%
Recall	0.49	0.68	+38.8%
Click-Through Rate (CTR)	1.74%	2.46%	+41.3%

Source: Author

Paired sample t-tests were used to examine differences in important performance factors. Results from these tests showed that there was a significant improvement in all three factors. On Precision, for example, the t-test gave a result of $t=6.83$, with $p < 0.001$, which means that a significant improvement occurred. In recall, a comparable improvement was observed, which had a value of $t=7.12$, with $p < 0.001$. Moreover, the improvement in CTR was found to be significant, with a value of $t=5.49$, with $p < 0.001$. Overall, it could be concluded that the chances of these observations occurring by chance were very low.

These data support the second hypothesis, indicating that cross-platform metadata integration is a significant positive predictor of targeting accuracy and engagement performance. In this respect, it is evident that the integration of behavioral metadata has lifted the targeting effectiveness of the system for delivering high-interest users into the campaign. Hence, metadata integration influences user engagement.

4.4. Engagement Lift

Engagement lift was analysed by comparing pre- and post-personalization interaction data for the same user segments. Engagement lift represents the depth and longevity of user interaction following personalization (Table 3). The observed increase in average session duration (from 1.9 to 3.4 minutes) and conversions (from 0.86% to 1.61%) reflects higher content relevance and reduced friction in the user journey.

Table 3: Engagement Metrics Before and After Metadata Integration.

Engagement Metric	Pre-Integration	Post-Integration	% Change
Average Session Duration	1.9 minutes	3.4 minutes	+78.9%
Pages per Session	2.7	4.3	+59.3%
Conversion Rate	0.86%	1.61%	+87.2%

Source: Author.

Such improvement was seen most strongly within Cluster 3 (Multi-Touch Converters), suggesting that personalization through metadata not only helps improve the accuracy of the first click but also helps users through the entire journey. The engagement seen here also highlights the commercial importance of personalization through cross-platform metadata.

Figure 3: Heatmap Showing Engagement Metrics Before And After Integration.

Source: Author.

From the data presented, it is clear that there has been an improvement in several key performance metrics, as depicted in Figure 3. For instance, Average Session Duration has increased from 1.9 to 3.4 minutes, an improvement of 78.9%. This shows an improvement in user engagement with the website. Additionally, Pages per Session has increased from 2.7 to 4.3, an improvement of 59.3%. This shows that there has been an improvement in user engagement, as users are accessing more pages on the website. Lastly, there has been an improvement in the Conversion Rate, as it has increased from 0.86% to 1.61%. This shows an improvement of 87.2%, which is an indication of an improvement in user engagement in activities aimed at fulfilling the desired objectives. From the analysis, two clusters, namely Cluster 3 (Multi-Touch Converters) and Cluster 2 (Engaged Explorers), are related to the improvement in user engagement. This shows the value of the unified customer profile in improving user engagement.

4.5. Chi-Square Analysis of Sectoral Variation

An assessment was carried out in order to evaluate if there existed a significant difference in the standardization of metadata with respect to a certain sector (see table 4), using a chi-square test with respect to the adoption of normalized schemas among five industry types. The results of the chi-square test show that there is an association between sector types and the adoption of metadata standardization. It is possible that certain sectors, such as telecommunications or retail, with more advanced digital infrastructures or real-time personalization, show a higher adoption rate of standardization.

Table 4: Metadata Standardization Across Sector.

Sector	Standardized Metadata (n)	Non-Standardized (n)
Retail	17	4
Telecom	16	1
Banking	13	2
Tourism	14	6
Education	15	8

Source: Author.

The value of the chi-square statistic is 11.64, and the degrees of freedom = 4, with a p-value of 0.020, showing a significant relationship between sector type and the adoption of metadata standardization techniques. These findings show the accuracy of the proposed conceptual framework, as it is plausible that the context of the respective industries moderates the adoption of metadata strategies, as constraints may be imposed on publicly funded or educational institutions.

4.6. Regression Analysis

A linear regression analysis was employed to predict personalization accuracy (measured by precision and recall) using metadata standardization score and integration completeness score as predictor variables. Analysis of the linear regression results (table 5) indicates that both metadata standardization and integration completeness constitute significant predictors of personalization accuracy. The regression analysis explains more than two-thirds of personalization success with an R-square of 0.67.

Table 5: Regression Model Predicting Personalization Accuracy.

Predictor	β Coefficient	t-Value	p-Value
Metadata Standardization	0.44	5.12	<0.01
Integration Completeness	0.52	6.73	<0.001
Model R^2	0.67		

Source: Author.

esnlztn ccucy

$$= 0 \quad + 1 \quad (tnztn)$$
$$+ 2 \quad (ntetn) +$$

The results of the analysis showed high explanatory power, as evident from the R^2 value of 0.67 for the model. This means that the model explains about 67% of the variability of the dependent variable. Both predictor standardization and integration were found to be significant. Standardization had a standardized beta weight (β_1)

of 0.44 with a significance level of $p < 0.01$, whereas integration was found to be slightly more important with a standardized beta weight (β_2) of 0.52 and a significance level of $p < 0.001$. These results clearly indicate the significant contribution of both variables to the model.

Both predictors were significant. Integration had a slightly more intense effect, indicating that in order for personalization to occur correctly, it is crucial that data is made unified between systems.

These results strongly support both hypotheses and indicate the business imperative of strategic efforts with respect to metadata for any data marketing campaign.

4.7. Summary of Results

These results from the process of clustering, performance analysis, tests of inference, and regression analysis cumulatively constitute proof for the validity of both hypotheses. The standardization of metadata resulted in a significant improvement in the integration of customer data, thus proving hypothesis one, and cross-platform integration resulted in the improvement of the accuracy of personalization and campaign engagement, thus proving hypothesis two. Further, chi-square analysis performed in various sectors showed that there are differences in metadata usage in a given context, thereby reiterating that industry moderates as per definitions of metadata frameworks. Results from all analyses support both of the study's hypotheses. The objective of this study, which aims to investigate metadata standardization and customer profiling, is met as per Objective 1 since clustering, compatibility of profiles, and regression analysis are successful. If we look at Objective 2, it focused upon evaluating the impact of integration on targeting precision and engagement, is confirmed by significant increases in accuracy metrics, session depth, and conversion performance.

4.8. Discussion

This paper endeavors to investigate how metadata standardization and cross-platform integration influence personalization accuracy in Saudi Arabia's electronic marketing environment. The discussion will focus on how the key concepts of metadata standardization, user profiling, cross-platform integration, and personalization accuracy relate to the topic in the frame of established empirical and theoretical knowledge. Rather than discussing statistical findings, will attempt to position findings in a broader knowledge domain. The literature has, of late, accepted that metadata has been a valuable factor in digital markets. Notable authors such as Indig et al. in 2022 and Mohammad et al. in 2025d, as well as Mohammad et al. in 2025 and Nitzani in 2019, cited metadata as being not only a supportive infrastructure but an enabler for content personalization, user journey analysis, and predictive analysis. Study contributes to this discourse as it explains that any value derived from metadata will be exponentially enhanced if it were standardized and integrated with all other different platforms, joining a comprehensive and dynamic user profile. This current study supports proposals of data fusion epistemology stipulated by Hassani et al. in 2023, where informational value derived from "user behavior mobilized from integrated data streams with consistent structure" has a value in itself that needs no cause for justification. The theoretical platform, as established in this present study for metadata standardization, has been adapted from schema standardization, which has been established in information systems and semantic web design, as cited in 2021 by authors Sawarkar and Kodati. The complementation of common names, common timing, and definition identifiers in different platforms has been established to provide a "semantic interoperability" as established in 2019 by authors Venceslau Venceslau, that integrated various data sources from different origins. This present study for metadata standardization has adapted from this model of standardization, which has been proven to enhance semantic interoperability for a far more effective use of electronic content. User profiling, with personalization, has long been established in classical models that were for segmented users and focused on individual demographics and transaction analysis, as established in 2024 among authors of "Janaki.". However, the evolution toward behaviour-based and event-driven targeting captured in real-time metadata has marked a paradigm shift. This represents a reflection of this shift and verifies the findings of Sawarkar and Kodati in 2021 that live segmentation informed by behavior has a higher predictive value in terms of user intent than demographic information. Moreover, using unsupervised machine learning in this study fits with similar work in the customer analytics literature, as found in Kristijansson and Aegisson in 2022, which argues that clustering analytics allow for reactions to user data that are proactive in nature.

This study is in line with Sun (2024) in dealing with the concerns and expectations that are projected in a fragmentation of customer data as a challenge for gaining meaningful insights. The fragmentation challenge that emerges from the data of the customers has been eased by Shiramshetty (2020), who emphasized data integration as a need for gaining better insights into customers. Variations in the type of sectors emerging in the application of metadata concepts in this research paper are linked to concepts of digital maturity models in literature (Ladu et al., 2024).

Sectors such as the telecommunication sector, which is proven to be relatively more adept at their digital strategies, are correctly positioned for the task of standardization of metadata. The educational sector, which are proven to be further constrained either by the inflexibilities of technology or by their own regulatory demands, are correctly positioned behind on the issue of adoption of this infrastructure. This clearly draws attention to a finding by a separate study by Pangrazio et al. (2022) who strongly posited that for effective data use, it is almost entirely the role of leaders and malleability of data structure. Concerning Saudi Arabia as the setting, the study draws key pointers to their findings in the context of a unique structural setting. Saudi Arabia is actually at a crossroads in terms of development of regulations and infrastructure development, distinguishing it from other more developed online markets in the global context. Possibly, cross-platform metadata integration is likely to fall short of such promise because of a lack of sound enforcement of the Personal Data Protection Law, lack of development of integrated data management systems. Additionally, disparities in the levels of development of various industries in general, focusing on industries such as public services, retail, and finance, make it difficult for the standardization of metadata systems or the development of a standardized profiling system.

The more widespread implications of such research connect with the discussion on data driven marketing ethics. Some authors, including Jeffers, (2020) guvxb344, have alerted about the implications of such tracking systems for surveillance purposes. It is important to state that this research paper refrains from engaging with any of the more qualitative issues related to data marketing ethics. Furthermore, the research data confirms its utility for real-time marketing, aligning with statements by the marketing industry that real-time personalization is an important step for boosted marketing results achieved through structurally accurate data systems (Rahman et al., 2024).

5. CONCLUSION

This study seeks to explore the influence of metadata standardization and cross-platform integration on the accuracy of personalization in digital marketing in the current Saudi market. Using cross-platform integration of metadata through platforms such as Facebook Insights with Google Analytics, among other regression analysis techniques, this study seeks to establish that the standardization and cross-platform integration of metadata contribute to enhanced efficiency in current online personalization. This study has several strengths; nevertheless, it also has certain weaknesses. Firstly, the study is limited to five industries in Saudi Arabia. This might restrict the applicability of the study to other markets that have different levels of digitalization and legal systems. Secondly, while the study uses different methodological approaches, it only uses aggregate consumer behaviour data without using tracking methods, which can affect predictive results. Finally, while the study discusses data privacy and ethical issues, it does not discuss the perceptions of the data by the users.

The above challenges can be addressed by future research, which can make cross-national comparisons to investigate the impact of data policy on the application of metadata integration. Additionally, longitudinal studies of user responses to personalization can make a significant contribution to the existing body of knowledge, especially with regard to exploring various aspects of personalization in metadata marketing. A more specific analysis of the application of various aspects of personalization with the aid of artificial intelligence, including real-time engines with deep learning algorithms, can make a significant contribution to the existing body of knowledge. The present study has successfully reaffirmed the significance of metadata integration for improving digital marketing outcomes. The present study has both theoretical and practical implications for digital marketing practitioners who aim to improve their application of personalization strategies in the complex digital marketing landscape.

Acknowledgement
The author extends his appreciation to the Deanship of Scientific Research, King Saud University for supporting this project.

REFERENCES

Ali, I., Patil, Y. V., Jangid, A., Rahaman, Md. A., Taru, R. D., & Iftikhar, A. (2025). Blockchain-Driven Supply Chain Finance for Public Healthcare in India: Enhancing Financial Resilience in Public Health Systems. Salud Ciencia y Tecnología, 5, 1400. https://doi.org/10.56294/saludcyt20251400

Amin, M. M. (2024). AI-Powered Personalized Marketing: A Deep Dive into Customer Segmentation and Targeting. International Journal of Science Engineering and Management, 11(12). https://doi.org/10.36647/ijsem/11.12.a001

Asha, A. I., -, S. K. H., -, M. A. I., -, S. afrin P., & -, N. M. I. (2024). The Role of Big Data Analytics in Personalized Marketing: Enhancing Consumer Engagement and Business Outcomes. International Journal For Multidisciplinary Research, 6(5). https://doi.org/10.36948/ijfmr.2024.v06i05.28077

Bailo, D., Paciello, R., Vinciarelli, V., Orfino, A., Jeffery, K., Michálek, J., & Goffi, C. (2023). Integrated Access to Multidisciplinary Data Through Semantically Interoperable Services in a Metadata-Driven Platform for Solid Earth Science. In Communications in computer and information science (p. 235). Springer Science+Business Media. https://doi.org/10.1007/978-3-031-39141-5_20

Barari, M., Ross, M., Quach, S., & Surachartkumtonkun, J. (2020). A meta-analysis of customer engagement behaviour. International Journal of Consumer Studies, 45(4), 457. https://doi.org/10.1111/ijcs.12609

Buranarach, M., Buranasing, W., Rungcharoensuksri, S., Sarawasee, P., Ngootip, T., & Chansanam, W. (2022). Metadata Integration Framework for Data Integration of Socio-Cultural Anthropology Digital Repositories: A Case Study of Princess Maha Chakri Sirindhorn Anthropology Centre. Informatics, 9(2), 38. https://doi.org/10.3390/informatics9020038

Chaturvedi, P. (2025). Personalization in Banking: The Key to Customer Retention. International Journal of Management and Humanities, 11(8), 14. https://doi.org/10.35940/ijmh.g1793.11080425

Chen, X., Mitra, S., & Swaminathan, V. (2020). Metadata Matters in User Engagement Prediction. 1529. https://doi.org/10.1145/3397271.3401201

Chester, J., & Montgomery, K. (2017). The role of digital marketing in political campaigns. Internet Policy Review, 6(4). https://doi.org/10.14763/2017.4.773

Darvidou, K. (2024). Omnichannel Marketing in the Digital Age: Creating Consistent, Personalized and Connected Customer Experiences. Technium Business and Management, 10, 34. https://doi.org/10.47577/business.v10i.11903

Dave, A., Kolli, R. K., Mokkapati, C., Goel, O., Khan, S., & Jain, Prof. A. (2022). Techniques for Enhancing User Engagement through Personalized Ads on Streaming Plat-forms. Universal Research Reports, 9(3), 196. https://doi.org/10.36676/urr.v9.i3.1390

Finn, G. (2021). Report: 89% Of Marketers Have Problems With Cross-Channel Marketing, Data Linkage Is The Biggest Culprit. https://martech.org/report-89-of-marketers-have-problems-with-cross-channel-marketing-data-linkage-is-the-biggest-culprit/

Gorgoglione, M., Palmisano, C., & Tuzhilin, A. (2006). Personalization in Context: Does Context Matter When Building Personalized Customer Models? Proceedings, 222. https://doi.org/10.1109/icdm.2006.125

Hassani, S., Dackermann, U., Mousavi, M., & Li, J. (2023). A systematic review of data fusion techniques for optimized structural health monitoring [Review of A systematic review of data fusion techniques for optimized structural health monitoring]. Information Fusion, 103, 102136. Elsevier BV. https://doi.org/10.1016/j.inffus.2023.102136

Hong, L., & Lalmas, M. (2019). Tutorial on Online User Engagement: Metrics and Optimization. 1303. https://doi.org/10.1145/3308560.3320087

Ho-Quang, T., Chaudron, M. R. V., Robles, G., & Herwanto, G. B. (2019). Towards an Infrastructure for Empirical Research into Software Architecture: Challenges and Directions. 34. https://doi.org/10.1109/ecase.2019.00014

Indig, B., Sárközi-Lindner, Z., & Nagy, M. (2022). Use the Metadata, Luke! – An Experimental Joint Metadata Search and N-gram Trend Viewer for Personal Web Archives. 47. https://doi.org/10.18653/v1/2022.nlp4dh-1.7

Iyer, C. (2019). Data collection for integrated customer data management: Concerns, considerations and best practices. Applied Marketing Analytics, 5(1), 38. https://doi.org/10.69554/sfad6235

Janaki, Mohan, N., Prabhu, M., Parveen, A., & Menaga, E. (2024). The Segmentation Revolution: Customer Personality Analysis Driving Predictive Success. 2022 8th International Conference on Advanced

Computing and Communication Systems (ICACCS), 1775. https://doi.org/10.1109/icaccs60874.2024.10716968

Jeffers, M. T. (2020). The privacy problem for internalizing behavioural externalities. Society and Economy, 43(1), 60. https://doi.org/10.1556/204.2020.00030

Kristijansson, A. M., & Aegisson, T. (2022). Survey on Technique and User Profiling in Unsupervised Machine Learning Method. Journal of Machine and Computing, 9. https://doi.org/10.53759/7669/jmc202202002

Ladu, L., Koch, C., Ashari, P. A., Blind, K., & Castka, P. (2024). Technology adoption and digital maturity in the conformity assessment industry: Empirical evidence from an international study. Technology in Society, 77, 102564. https://doi.org/10.1016/j.techsoc.2024.102564

Li, Z., Sun, W., Zhan, D., Kang, Y., Chen, L. Y., Bozzon, A., & Hai, R. (2024). Amalur: Data Integration Meets Machine Learning. IEEE Transactions on Knowledge and Data Engineering, 1. https://doi.org/10.1109/tkde.2024.3357389

Matz, S., & Kosiński, M. (2019). Using Consumers' Digital Footprints for More Persuasive Mass Communication. NIM Marketing Intelligence Review, 11(2), 18. https://doi.org/10.2478/nimmir-2019-0011

Mohammad, A. A. S. (2025). The impact of COVID-19 on digital marketing and marketing philosophy: evidence from Jordan. International Journal of Business Information Systems, 48(2), 267-281.

Mohammad, A. A. S., Al-Hawary, S. I. S., Oraini, B. A., Vasudevan, A., Hindieh, A., Altarawneh, A., Alshurideh, M. T., & Ali, I. (2025). Strategies for applying interpretable and explainable AI in real world IoT applications. Discover Internet of Things, 5(1). https://doi.org/10.1007/s43926-025-00155-z

Mohammad, A. A. S., Mohammad, S. I. S., Al Oraini, B., Vasudevan, A., & Alshurideh, M. T. (2025c). Data security in digital accounting: A logistic regression analysis of risk factors. International Journal of Innovative Research and Scientific Studies, 8(1), 2699-2709.

Mohammad, A. A. S., Mohammad, S. I. S., Al-Daoud, K. I., Al Oraini, B., Vasudevan, A., & Feng, Z. (2025b). Optimizing the Value Chain for Perishable Agricultural Commodities: A Strategic Approach for Jordan. Research on World Agricultural Economy, 6(1), 465-478.

Mohammad, A. A. S., Mohammad, S. I. S., Al-Daoud, K. I., Vasudevan, A., & Hunitie, M. F. A. (2025d). Digital ledger technology: A factor analysis of financial data management practices in the age of blockchain in Jordan. International Journal of Innovative Research and Scientific Studies, 8(2), 2567-2577.

Mohammad, A. A., Shelash, S. I., Saber, T. I., Vasudevan, A., Darwazeh, N. R., & Almajali, R. (2025a). Internal audit governance factors and their effect on the risk-based auditing adoption of commercial banks in Jordan. Data and Metadata, 4, 464

Nitzani, M. (2019). Knowing You Better than You Know Yourself. In Routledge eBooks (p. 204). Informa. https://doi.org/10.4324/9780429280108-17

Odedina, C. (2023). Impact of Big Data on Marketing Strategy and Consumer Behaviour Analysis in the Us. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.4520361

Osakwe, J., Shilongo, A., & Ziezo, M. M. (2023). Optimising Customer Segmentation in Digital Marketing Using Predictive Analytics: A Review of Literature [Review of Optimising Customer Segmentation in Digital Marketing Using Predictive Analytics: A Review of Literature]. SSRN Electronic Journal. RELX Group (Netherlands). https://doi.org/10.2139/ssrn.4662191

Pangrazio, L., Selwyn, N., & Cumbo, B. (2022). A patchwork of platforms: mapping data infrastructures in schools. Learning Media and Technology, 48(1), 65. https://doi.org/10.1080/17439884.2022.2035395

Paul, J. (2019). Marketing in emerging markets: a review, theoretical synthesis and extension. International Journal of Emerging Markets, 15(3), 446. https://doi.org/10.1108/ijoem-04-2017-0130

Pesquita, C., Ivanova, V., Lohmann, S., & Lambrix, P. (2018). A Framework to Conduct and Report on Empirical User Studies in Semantic Web Contexts. In Lecture notes in computer science (p. 567). Springer Science+Business Media. https://doi.org/10.1007/978-3-030-03667-6_36

Rahman, A. V., Dika, R. P., & Khairi, A. (2024). Personalization vs. Privacy: Balancing Customer Data Use with Ethical Marketing Practices. Equator Journal of Management and Entrepreneurship (EJME), 12(4), 349. https://doi.org/10.26418/ejme.v12i4.82521

Rana, N. P., Slade, E., Sahu, G. P., Kizgin, H., Singh, N., Dey, B. L., Gutiérrez, A., & Dwivedi, Y. K. (2019). Digital and Social Media Marketing. In Advances in theory and practice of emerging markets.

Springer International Publishing. https://doi.org/10.1007/978-3-030-24374-6

SadighZadeh, S., & Kaedi, M. (2022). Modeling user preferences in online stores based on user mouse behaviour on page elements. Journal of Systems and Information Technology, 24(2), 112. https://doi.org/10.1108/jsit-12-2019-0264

Sarfraz, N., Syed, T., & Khan, U. A. (2021). A STUDY ON PERCEPTIONS OF MARKETING PROFESSIONALS TOWARDS THE DIGITAL MARKETING IN SAUDI ARABIA. Journal of Social Research Development, 2(2), 120. https://doi.org/10.53664/02-02-2021-02-120-128

Sawarkar, K., & Kodati, M. (2021). Automated Metadata Harmonization Using Entity Resolution and Contextual Embedding. In Lecture notes in networks and systems (p. 129). Springer International Publishing. https://doi.org/10.1007/978-3-030-80126-7_10

Seaver, N. (2021). Seeing like an infrastructure: avidity and difference in algorithmic recommendation. Cultural Studies, 35, 771. https://doi.org/10.1080/09502386.2021.1895248

Shiramshetty, S. K. (2020). Data Integration Techniques for Cross-Platform Analytics. International Journal of Scientific Research in Computer Science Engineering and Information Technology, 593. https://doi.org/10.32628/cseit2064139

Singh, N., Sunkara, R., & Yencken, S. (2021). Power, Privacy and Personalization in Digital Commerce. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.3780726

Slanzi, G., Pizarro, G., & Velásquez, J. D. (2017). Biometric information fusion for web user navigation and preferences analysis: An overview. Information Fusion, 38, 12. https://doi.org/10.1016/j.inffus.2017.02.006

Sun, C. (2024). Data Analysis of Customer Segmentation and Personalized Strategy in the Era of Big Data. Advances in Economics Management and Political Sciences, 92(1), 46. https://doi.org/10.54254/2754-1169/92/20231411

Tarifi, N., & Bakhsh, R. (2024). The Effectiveness of Personalized Marketing Strategies on Consumer Engagement and Loyalty (Evidence from KSA). International Journal of Research and Studies Publishing, 5(56), 283. https://doi.org/10.52133/ijrsp.v5.56.11

Tu, H., Lin, Z., & Lee, K. (2019). Automation With Intelligence in Drug Research. Clinical Therapeutics, 41(11), 2436. https://doi.org/10.1016/j.clinthera.2019.09.002

Udantha, M., Ranathunga, S., & Dias, G. (2016). Modelling website user behaviors by combining the EM and DBSCAN algorithms. 2022 Moratuwa Engineering Research Conference (MERCon), 168. https://doi.org/10.1109/mercon.2016.7480134

Venceslau, A., Andrade, R. M. C., Vidal, V. M. P., Nogueira, T. P., & Pequeno, V. M. (2019). IoT Semantic Interoperability: A Systematic Mapping Study. https://doi.org/10.5220/0007732605350544

Waehner, K. (2024). Unified Commerce in Retail and eCommerce with Apache Kafka and Flink for Real-Time Customer 360. https://www.kai-waehner.de/blog/2024/08/30/unified-commerce-in-retail-and-ecommerce-with-apache-kafka-and-flink-for-real-time-customer-360/

Yoneda, T., Kozawa, S., Osone, K., Koide, Y., Abe, Y., & Seki, Y. (2019). Algorithms and System Architecture for Immediate Personalized News Recommendations. IEEE/WIC/ACM International Conference on Web Intelligence, 2651, 124. https://doi.org/10.1145/3350546.3352509