Article

Not peer-reviewed version

Information Sustainability Beyond Digital Access: Evidence of Machine Learning in Local Media Ecosystems in Ecuador

Luis Saráuz-Estevez *, Jessica Pupiales-Proaño, Danilo Cuaical-Tapia

Posted Date: 13 March 2026

doi: 10.20944/preprints202603.1050.v1

Keywords: knowledge management; information sustainability; digital transformation; digital divide; media consumption profiles; local media ecosystems; machine learning; cluster analysis; random forest; Ecuador

Article

Information Sustainability Beyond Digital Access: Evidence of Machine Learning in Local Media Ecosystems in Ecuador

Luis Saráuz-Estevez *, Jessica Pupiales-Proaño and Danilo Cuaical-Tapia

Facultad de Ciencias Administrativas y Económicas, Universidad Técnica del Norte, Ibarra 100150, Ecuador

* Correspondence: lvsarauz@utn.edu.ec (L.S-E.); Tel.: +593-0996645642

Abstract

Information sustainability has emerged as a key dimension of social sustainability, as it supports equitable access to information, informed opinion formation, and institutional engagement. However, the expansion of digital platforms does not necessarily guarantee inclusive information ecosystems, particularly in local contexts characterized by structural inequalities. This study examines information sustainability and the digital divide by identifying media consumption profiles within a territorial context in Ecuador. Using data from a survey conducted in the province of Imbabura with 1,784 observations, a hybrid methodological approach combining cluster analysis and Random Forest (RF) algorithms was applied. Audience profiles were identified and validated based on media consumption patterns, levels of digitalization, and institutional engagement. The results reveal four distinct audience profiles with different levels of digital integration and institutional linkage. Findings indicate that the intensity and diversity of media consumption play a more decisive role than mere technological access. Digital access alone is insufficient to ensure information sustainability or foster institutional opinion formation; instead, differences in exposure, usage intensity, and media habits shape audience engagement. These findings highlight the need for segmented, territory-based communication strategies to strengthen information sustainability, reduce the digital divide, and reinforce the role of university media within local media ecosystems.

Keywords: knowledge management; information sustainability; digital transformation; digital divide; media consumption profiles; local media ecosystems; machine learning; cluster analysis; random forest; Ecuador

---

1. Introduction

Sustainability is a multifaceted concept that transcends strictly environmental approaches and incorporates social, informational and institutional dimensions. From the perspective of social sustainability, the emphasis is on the impact that organisations have on people and society, with the aim of ensuring equitable access to social resources, opportunities and quality of life for present and future generations [1]. In this context, information takes on a central role, as it is an essential resource for citizen participation, social cohesion and informed decision-making.

Improving well-being, ensuring fair distribution of resources, valuing diversity, fostering social cohesion and promoting inclusive governance are key principles of social sustainability. In various productive sectors, such as textiles and agriculture, these principles translate into practices that have a positive impact on workers, local communities and society as a whole [2]. However, in the field of information, these objectives are conditioned by the profound transformations resulting from the digitalisation of media ecosystems and by persistent inequalities in access to and use of information.

Although the expansion of digital platforms has broadened opportunities for access to information and interaction between institutions and audiences, this process does not automatically guarantee greater institutional participation or better public opinion formation. The literature warns

 doi:10.20944/preprints202603.1050.v1

that digitisation can have ambivalent effects: while excessive institutional influence can lead to polarisation, a moderate and sustained presence has greater potential to foster consensus, trust and institutional legitimacy [3]. Consequently, digital access alone is not a sufficient condition for ensuring information sustainability or institutional commitment.

In this regard, the digital disconnect observed in certain segments of the population does not necessarily reflect attitudes of rejection towards information or institutions, but rather, in many cases, is due to structural limitations associated with unequal access to devices, connectivity and digital skills. This situation is particularly evident in economically disadvantaged communities, where such deficiencies restrict full participation in contemporary information environments [4–6]. These inequalities not only affect access to digital education [7], but also limit the possibility of forming an informed opinion and establishing sustained links with institutions.

In the field of education, university media outlets are strategic players in social and informational sustainability, insofar as they contribute to disseminating knowledge, the strengthening of institutional ties, and the formation of opinion in the territories where they operate. However, these media face the challenge of adapting to new information consumption habits, particularly among young audiences, which involves exploring short digital formats, social media-oriented content, and participatory strategies that involve students themselves in media production [8]. The ability of university media to fulfil their social function therefore depends on their effective articulation with local and digital media ecosystems.

In Ecuador, empirical evidence shows a predominant use of social media as a source of information, accompanied by greater risks of misinformation and significant differences in media literacy, content verification, and trust in information sources [9]. Similarly, gaps have been identified in relation to access, use and information skills, with direct effects on decision-making and the performance of different social groups [10,11], as well as the central role of social media in shaping political imaginaries and public opinion, especially among young people [12].

By contrast, international and national studies on institutional communication have focused mainly on analysing content strategies and the use of social media by institutions, revealing predominantly informative and one-way communication, with low levels of segmentation and personalisation [13–16]. Furthermore, research examining institutional perception, credibility, or trust continues to rely heavily on traditional statistical methods, such as regressions or structural equation models, without systematically incorporating machine learning techniques [17–19].

There are still few studies that apply machine learning approaches to segment audiences and analyse media consumption patterns in relation to institutional perception and engagement. Most of this research focuses on content personalisation or general characterisation of digital behaviour, without integrating media consumption, the level of digitalisation and the construction of institutional opinion [20–23]. In Ecuador, machine learning-based studies focus on predicting digital behaviour, without comprehensively addressing these dimensions [24], which highlights a particularly relevant knowledge gap in local Latin American contexts.

This study contributes to the literature by operationalising information sustainability through empirical segmentation based on machine learning, expanding predominantly normative approaches. Furthermore, Latin America presents hybrid media dynamics that combine analogue tradition and uneven digitisation, making the current Ecuadorian case a relevant laboratory for the analysis of information sustainability.

In response to this gap, this article aims to identify audience profiles based on media consumption and level of digitalisation, and to analyse how these profiles relate to the construction of institutional opinion in an Ecuadorian context. To this end, the following hypotheses are formulated:

H1. The level of digital media usage significantly influences the configuration of differentiated information consumption profiles within the local media ecosystem.

H2. A higher level of information digitisation does not guarantee greater institutional engagement, but rather gives rise to hybrid configurations of media consumption.

H3. The frequency of consumption of institutional media is positively associated with a favourable assessment of academic management and university content.

H4. Information profiles in the local media ecosystem present hybrid configurations characterised by a combination of traditional, digital and institutional media consumption at varying levels of intensity.

H5. The digital divide in local contexts is a multidimensional phenomenon that is better explained by differences in media usage patterns and institutional valuation than by isolated technological access.

H6. Variables related to digitisation and institutional consumption enable robust predictions to be made about membership of different information profiles using machine learning models.

To this end, a methodology based on cluster analysis and Random Forest models is applied to generate empirical evidence. The findings are intended to support the design of segmented institutional communication strategies aimed at strengthening information sustainability and reducing digital divides in local media ecosystems.

2. Literature Review

2.1. Social and Informational Sustainability

Equitable access to information is an essential component of social sustainability, as it ensures that people can form informed opinions and participate actively and responsibly in social, political and institutional life. This approach is closely linked to the principles of social justice, equity and community cohesion, and is particularly relevant in contexts where information inequalities disproportionately affect socially and economically vulnerable groups [25,26].

From a democratic perspective, access to diverse and reliable information is essential for the exercise of citizenship, as it allows individuals to critically evaluate public management, demand accountability, and make informed decisions [27,28]. However, this ideal is strained by persistent structural inequalities and the influence of power relations that can limit the visibility of certain voices and perspectives within the information ecosystem [29].

The notion of information sustainability (IS) is linked to the role of information systems and media ecosystems in supporting the Sustainable Development Goals (SDGs). Several studies indicate that sustainable management of information systems can contribute to improving resource efficiency, strengthening social responsibility, and generating long-term economic and social value [30,31]. This mutual action of sustainability in IS management involves assessing the environmental, social and economic impacts of these systems [32]. The integration of sustainability criteria into information management therefore involves jointly assessing the social, economic and organisational impacts of communication systems and channels.

The formation of informed opinion is thus a central pillar of social sustainability, insofar as it directly affects equity, social cohesion, and community participation. Ensuring equitable information conditions not only promotes collective well-being, but also strengthens social resilience and the capacity of communities to face present and future structural challenges [1,33].

2.2. The Digital Divide and Media Use

The digital divide refers to inequalities in access to, use of, and benefit from digital technologies and information content. This phenomenon is not limited to a strictly technological dimension, but reflects deep social inequalities associated with factors such as socioeconomic status, education, age, gender, and geographical location [34,35].

The literature distinguishes between different levels of the digital divide. The first level relates to physical access to technological infrastructure, including the availability of devices and internet connectivity, where economic and territorial conditions play a decisive role. Therefore, improving infrastructure, especially in rural or peripheral areas, is key to reducing this type of exclusion [36].

 doi:10.20944/preprints202603.1050.v1

The second level of the digital divide is linked to differences in digital skills and usage patterns. Previous studies show that higher levels of education and income are associated with more diverse and sophisticated use of digital technologies, while a lack of digital literacy limits meaningful engagement with information environments [37–39].

For its part, the third-level digital divide focuses on the results derived from the use of digital technologies, such as economic, social and civic benefits. Empirical evidence shows that the impact of internet use is not uniform and that, in certain contexts, groups with fewer resources may benefit unequally, reinforcing the need for public policies and institutional strategies aimed at ensuring equitable outcomes [36,40].

In the media sphere, the maturity of digital media varies significantly between countries and regions, conditioned by economic, infrastructural and political factors [41]. Likewise, the use of social media can contribute both to reducing and widening digital divides, depending on patterns of access, use and content consumed [42,43]. In contrast, the consumption of traditional media continues to be relevant, although the transition to digital environments has highlighted the importance of digital skills for meaningful information use, which can intensify existing knowledge gaps [44,45].

Overall, the digital divide is a multifaceted phenomenon that directly affects access to information and media consumption habits. Addressing it requires a comprehensive approach that combines improvements in physical access, the development of digital skills, and the promotion of equitable information outcomes.

2.3. Media Consumption and Opinion Formation

Media consumption plays a decisive role in shaping public opinion, as it influences the way individuals process, interpret and react to available information. Forming informed opinions helps to strengthen community participation, promote equity and encourage more effective decision-making processes, which are key elements in the development of more resilient and sustainable societies [46].

Various sociodemographic factors influence media consumption patterns and, consequently, opinion formation. Age, educational level and political interest condition the choice of information sources, so that young people with higher levels of education tend to favour digital environments, while older groups show a greater preference for traditional media such as television [47,48]. In a political context, media consumption can generate polarised opinions. In China, for example, state media and social media platforms have a different impact on public opinion on foreign policy [49].

From a dynamic perspective, theoretical and mathematical models have shown that the media can accelerate the spread of opinions and influence the balance of public opinion, especially when information interventions are directed at individuals with greater attitudinal flexibility [50,51]. Likewise, the influence of opinion leaders, media celebrities, and influencers, through parasocial relationships, can significantly shape public perceptions and attitudes [52].

In this sense, media consumption, whether through digital platforms or traditional media, not only facilitates the circulation of information, but also conditions the possibility of developing an institutionally informed opinion. The absence of an explicit opinion can therefore be interpreted as an indicator of low information exposure rather than a negative or dismissive attitude, which is particularly relevant for the analysis of audiences with limited media exposure [53,54].

2.4. Universities and Institutional Media

Universities play a strategic role in promoting social and informational sustainability, as they act as spaces for the production, dissemination, and legitimisation of knowledge. Through their institutional media and digital platforms, higher education institutions contribute to the formation of an informed citizenry and the strengthening of ties with the territories where they operate [55,56].

The integration of traditional, digital and university media creates a media ecosystem with the potential to drive local development and community participation. While traditional media remain relevant, their integration with digital platforms allows for the creation of more interactive and transparent information environments, expanding opportunities for learning and dialogue with audiences [57].

Likewise, incorporating sustainability into the core functions of universities, such as teaching, research and engagement with society, helps to promote long-term social change [58–60]. The way in which universities communicate their sustainability initiatives and values has a direct impact on their institutional image and public perception, reinforcing their position as key players in sustainable development [13,61].

However, recent studies warn that, despite the potential of social media and institutional media, levels of interaction with audiences tend to be limited, which restricts their capacity for dialogue and their informative impact [58,62]. In this scenario, the strategic and segmented use of university media is essential to strengthen information sustainability, foster institutional commitment, and reduce gaps in access and participation in local media ecosystems [63–65].

2.5. Theoretical Basis

Media Ecosystem Theory constitutes the conceptual framework underpinning this research. This perspective, initially developed by McLuhan [66] and Postman [67] and later expanded in the digital context by Napoli [68], Nambisan et al. [69] and Verhoef et al. [70]. It argues that media do not operate in isolation, but as interdependent systems where technologies, institutions and audiences co-evolve dynamically.

From this perspective, the media environment functions as a complex ecosystem in which each element, such as digital platforms, traditional media, institutions, and users, structurally influences the configuration of information behaviours. Digital transformation does not imply the linear replacement of traditional media with digital media, but rather a systemic reconfiguration of the communication environment, where hybrid patterns of consumption, trust, and institutional ties emerge.

Applied to the university and territorial context, this theory allows us to understand that information sustainability does not depend exclusively on technological access, but rather on the interaction between:

- Intensity of media consumption
- Level of digitisation
- Institutional perception
- Content assessment

3. Methodology

This study adopts a quantitative, explanatory and predictive approach, combining unsupervised multivariate analysis techniques and supervised machine learning, with the aim of identifying, structuring and validating types of media consumption and institutional interaction within the university communication ecosystem.

Specifically, Multiple Correspondence Analysis and Hierarchical Clustering on Principal Components (MCA+HCPC) with Random Forest (RF) were combined following recent recommendations in applied organisational sustainability research, which call for pairing structural segmentation with predictive validation to strengthen methodological robustness [71–73].

3.1. Database

The information gathering tool consisted of a structured survey designed to comprehensively measure computer usage patterns and levels of institutional interaction in the province of Imbabura, with a total of 1,784 observations.

The questionnaire was developed using a multidimensional approach, considering both exposure to traditional media and the transition to digital environments, as well as the institutional perception associated with university media. This structure allowed us to capture not only the frequency of use but also the intensity, preference, satisfaction, and ranking of alternative platforms.

The instrument consisted of 34 items organised into six clearly differentiated thematic blocks, which are presented in relation to the underlying theory in the following table:

Table 1. Relationship between the underlying theory and the sections of the questionnaire used in the survey.

Questionnaire Block	Component of the Media Ecosystem (Basic Theory)	Analytical Dimension	Operational Variables	Role in the Model
1. Traditional media (radio and television)	Traditional media subsystem	Conventional media exposure	Radio and television consumption frequency (CF_MT), Channel preference	Secondary structural dimension (MCA)
2. Print media	Persistence of the offline ecosystem	Print media consumption	Reading and press acquisition (CF_MI)	Complementary variable in MCA
3. Digital media (reading and use of platforms)	Digitisation of the media ecosystem	Intensity and digital transition	Platform usage (UP_DI), Digital frequency (F_U), Digital press readership (CF_MD)	Main axis of differentiation (MCA and RF – Hierarchical level 1)
4. Institutional opinion	Institutional integration within the ecosystem	Symbolic capital and institutional trust	Institutional consumption (CF_U), UTV assessment (SC_UTV), UTN management assessment (ER_UTN)	Institutional structuring dimension (MCA and RF – Hierarchical level 2)
5.Sociodemographic data	Structural determinants of the social system	Contextual factors	Age, gender, occupation, level of education, territory	Control variables and cluster characterisation
6. Internal field control	Structural validation of the measurement system	Methodological quality	Monitoring and validation variables	Guarantee of survey consistency

The instrument mainly integrated qualitative variables of a nominal and ordinal nature, including nominal polytomous categories associated with media preference, occupation and geographical location, as well as Likert-type ordinal scales to measure frequency of consumption and intensity of satisfaction.

 doi:10.20944/preprints202603.1050.v1

Due to the nature of the instrument variables, which are categorical and ordinal Likert-type, this study used MCA as an appropriate technique for modelling latent structures in categorical data without imposing assumptions of normality [74,75]. The use of this technique is particularly appropriate in sustainable segmentation studies, where perception and consumption patterns are often structured in unobservable latent dimensions [76–78].

3.2. Dimensionality Reduction Using Multiple Correspondence Analysis (MCA)

Multiple Correspondence Analysis (MCA) was used as a dimension reduction technique, which allows latent structures to be modelled in high-dimensional contingency matrices without requiring assumptions of normality or linearity, making it a suitable tool for studies of institutional behaviour and perception [77], since the information collection instrument was composed mainly of ordinal categorical variables.

In contexts of organisational sustainability and institutional communication, where constructs are often composed of perceptions, frequencies of use, and qualitative assessments, the MCA facilitates identifying underlying structural axes that organise collective behaviour [79].

Table 2. Results of Multiple Correspondence Analysis (MCA).

Indicator	Dimension 1	Dimension 2
Intrinsic value	0.264	0.172
Proportion of inertia explained	60.5%	39.5%

The MCA results (Table 2) show that the first two dimensions account for all the inertia retained in the reduced factorial space, confirming the existence of a stable two-dimensional structure that is sufficiently informative for segmentation purposes.

In methodological terms, Dimension 1 had an eigenvalue of 0.264 and explained 60.5% of the total retained inertia, constituting the main axis for structuring the data space. The operational interpretation of this axis was based on the categories with the highest absolute coordinates, observing a concentration of modalities associated with institutional consumption frequency (CF_U), content evaluation (CF_E) and perception of UTV (SC_UTV). Consequently, this axis was methodologically defined as a gradient of intensity and assessment of the institutional link.

Dimension 2, meanwhile, had a value of 0.172 and explained 39.5% of the retained inertia. The categories with the greatest geometric weight on this axis were linked to the use of digital platforms (UP_DI) and frequency of digital use (F_U), which allowed it to be interpreted operationally as an axis of digitisation and information patterns.

It is important to note that this interpretation is carried out exclusively for structural purposes, to understand the geometric organisation of the factorial space prior to hierarchical classification. The dimensions of the MCA represent latent axes of structuring of information behaviour and not direct observable variables, in line with the theoretical foundations of geometric data analysis.

3.3. Hierarchical Clustering on Principal Components (HCPC)

In order to identify homogeneous profiles of information behaviour, the Hierarchical Clustering on Principal Components (HCPC) classification procedure was applied, which combines factor reduction and hierarchical clustering in a sequential strategy. This approach has proven particularly suitable in segmentation studies based on categorical variables, as it allows for reducing the dimensionality of the data space before applying clustering techniques, improving the stability and interpretability of the clusters [80].

The classification was performed on the factorial coordinates obtained from the MCA, using Euclidean distance as a proximity metric. Subsequently, Ward's method was applied, which

 doi:10.20944/preprints202603.1050.v1

minimises intra-cluster variance at each fusion stage, favouring the formation of compact and well-differentiated groups [81]. This criterion is widely recommended in recent research in social sciences and sustainability due to its robustness in the face of complex categorical structures [82,83]. The number of clusters was determined based on inspection of the dendrogram and analysis of changes in inertia between consecutive levels of partitioning, complementing statistical criteria with theoretical consistency.

3.4. Supervised Validation Using Random Forest

In order to evaluate the structural consistency of the clusters identified using HCPC, a supervised Random Forest (RF) model was implemented. This procedure made it possible to compare the predictive stability of the clusters and reduce the exclusive reliance on unsupervised techniques.

The Random Forest algorithm, proposed by Breiman [84], is an ensemble learning method based on building multiple decision trees on bootstrap subsets of the data and random selection of variables at each node. This approach reduces the variance of the individual model and improves generalisation capacity, and is widely used in multi-class classification contexts with complex structures [85].

In the present study, the clusters obtained through HCPC were used as the target variable, while the original categorical variables constituted the set of predictors. This design allowed us to evaluate whether the identified segmentation structure can be reproduced through a supervised model, which constitutes an indirect form of internal validation.

3.5. Hyperparameter Tuning

The model was calibrated using grid search with five-fold stratified cross-validation. Different configurations of number of trees, maximum depth, and number of variables considered in each division were explored, selecting the optimal combination based on the Macro-F1 score. This metric was chosen due to its suitability in multi-class contexts with potentially unequal group sizes, as it equally weights performance in each category and avoids biases derived from the domination of majority classes.

3.6. Performance Appraisal

The final performance of the model was evaluated both on an independent test set (hold-out) and through cross-validation (out-of-fold) on the entire dataset. Multiple complementary metrics were reported: accuracy, balanced accuracy, Macro-F1, and Kappa coefficient. This combination provides a comprehensive assessment of predictive performance and reduce dependence on a single global indicator, in line with recent recommendations in explanatory modelling based on ensembles [86].

3.7. Robustness and Replicability

Dimensional reduction using MCA allowed the categorical space to be structured into interpretable latent dimensions, reducing redundancies and favouring more consistent segmentation. Subsequently, supervised validation using Random Forest was implemented with systematic hyperparameter search. The robustness of the model was evaluated using cross-validation procedures to ensure the stability and generalisation of the results.

Multiple evaluation metrics (accuracy, balanced accuracy, Macro-F1 and Kappa coefficient) were used to avoid dependence on a single global indicator. The analysis of variable importance, both by impurity reduction and permutation importance, allowed us to verify the consistency between the unsupervised structure and the supervised classification. Finally, the use of fixed seeds in stochastic algorithms ensures computational replicability, reinforcing the transparency and traceability of the analytical process.

 doi:10.20944/preprints202603.1050.v1

4. Results

4.1. Factor Structure of the Information Space

Multiple Correspondence Analysis allowed us to identify a stable two-dimensional structure that organised information behaviour into two main axes. The first dimension explained 60.5% of the retained inertia, while the second dimension explained 39.5%, together concentrating all the variability considered in the reduced space.

Dimension 1 showed a gradient associated with the intensity and assessment of the institutional link. On the positive side were categories related to higher frequency of institutional consumption (CF_U), favourable assessments of content (CF_E) and structured opinions on UTV (SC_UTV). In contrast, the negative side concentrated categories linked to low consumption, institutional ignorance or less favourable assessments. This configuration suggests that the main axis differentiates levels of institutional involvement.

Dimension 2 distinguished patterns of information digitisation. The categories with the highest coordinates associated with the use of digital platforms (UP_DI) and frequency of digital consumption (F_U), forming an axis that differentiates profiles with high dependence on digital media from less digitised information schemes.

4.2. Identification of Profiles Using HCPC

The application of the HCPC procedure on the factorial coordinates allowed four distinct clusters to be identified.

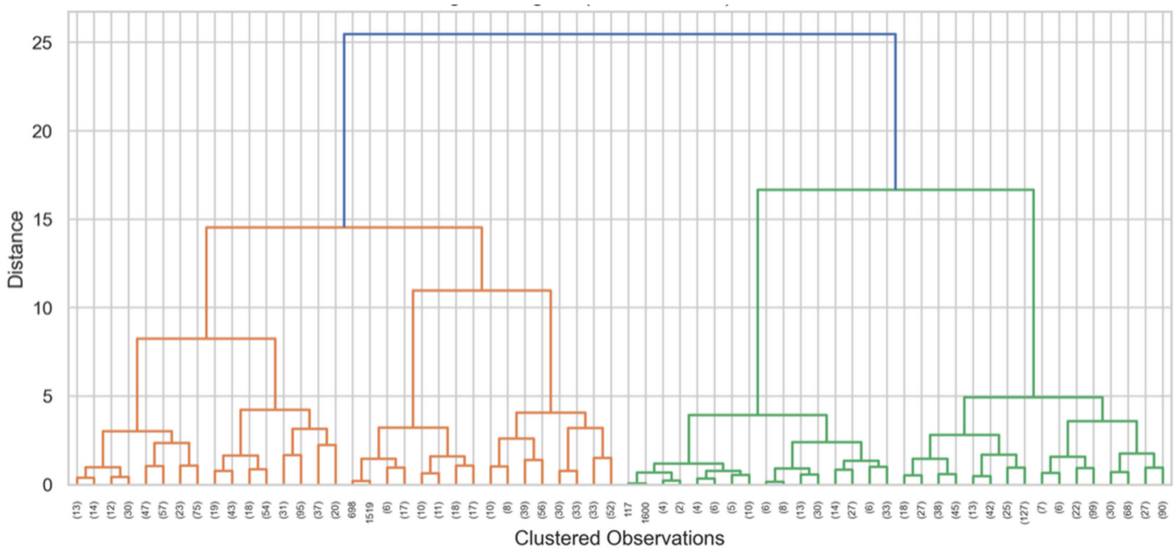

Figure 1. Dendrogram—Ward's method on MCA dimensions.

Table 3. *Final distribution of HCPC observations.*

HCPC cluster distribution	Cases
Cluster 1	342
Cluster 2	588
Cluster 3	170
Cluster 4	684

The analysis of typified residues allowed each group to be characterised based on significantly overrepresented categories (|z|≥2). The profiles identified can be summarised as follows:

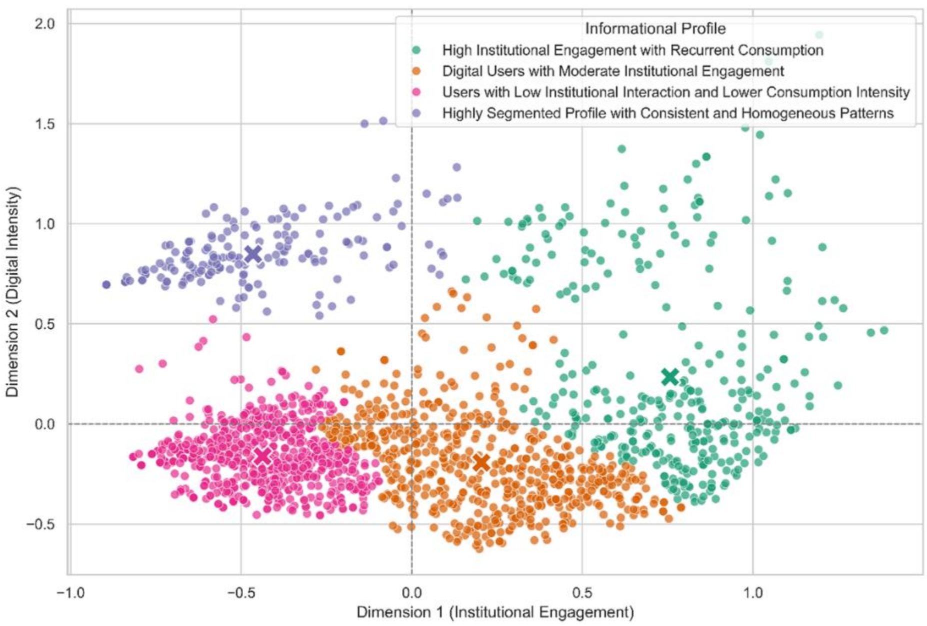

Figure 2. Informational Profiles in the MCA Space.

The profiles that reflect different combinations of institutional digitisation, consistent with the previously identified factor structure, are presented below:

- Cluster 1: High institutional engagement with recurring consumption

This profile is characterised by a high frequency of institutional consumption (CF_U), favourable content evaluations (CF_E) and structured opinions on UTV (SC_UTV). Individuals belonging to this group have a consolidated link with the institutional ecosystem, combining frequent interaction with positive assessment. From a structural point of view, this cluster is located at the positive pole of the institutional intensity axis, indicating a high level of involvement. It is not just about consumption, but a consistent pattern of recognition and favourable evaluation.

- Cluster 2: Digital users with moderate institutional ties

This group combines active use of digital platforms (UP_DI) and medium-high frequency of digital consumption (F_U) with moderate institutional engagement. Individuals with this profile obtain information mainly through digital channels, but their relationship with UTV and UTN does not reach the levels observed in cluster 1. This is a digital-first profile with functional rather than identity-based institutional interaction. Digitalisation does not imply disengagement, but rather a less intense and more instrumental relationship with the institutional environment.

- Cluster 3: Highly segmented profile with consistent, homogeneous patterns

The third cluster presents consistent and highly differentiated patterns, which is reflected in its high predictive separability in the supervised model. This profile combines categories that form a coherent internal structure, suggesting well-defined and relatively stable informative behaviours.

- Cluster 4: Users with low levels of institutional interaction and lower consumption intensity

This group concentrates categories associated with lower consumption frequency, lower institutional evaluation and lower interaction intensity. It is located at the negative pole of the institutional linkage axis, showing a pattern of low integration with the media ecosystem analysed. This profile does not necessarily imply a total absence of digital consumption, but rather lower intensity and less structuring of the institutional link, forming a group with limited interaction and less involvement.

4.3. Predictive Validation of Clusters

 doi:10.20944/preprints202603.1050.v1

The Random Forest model showed robust performance in both cross-validation and independent test sets. This algorithm was selected for its ability to model non-linear relationships and capture complex interactions between categorical variables. It also reduces the risk of overfitting through bootstrap aggregation, characteristics documented in multi-class classification applications in social sciences and sustainability studies [87,89].

Hyperparameter optimisation was performed using GridSearchCV with five-fold stratified cross-validation, using Macro-F1 as the target metric. The use of this metric is justified in multi-class problems with unequal group sizes, where Accuracy can overestimate performance in majority classes [90]. The optimal model selected 300 trees (n_estimators = 300), max_features = sqrt and free depth (max_depth = None), a configuration that showed performance convergence and adequate generalisation capacity.

Table 4. *Performance of the Random Forest model.*

Metric	Hold-out (TEST)	OOF (Cross-validation)
Accuracy	0.922	0.932
Balanced Accuracy	0.927	0.934
Macro-F1	0.932	0.939
Cohen's Kappa	0.887	0.903

The Random Forest model showed robust performance in both the independent test set and cross-validation (Table 4). The hold-out set obtained an accuracy of 0.92, a balanced accuracy of 0.93, and a Macro-F1 of 0.93, with a Kappa coefficient of 0.89, indicating high agreement beyond chance. These results suggest that the segmentation presents clear discriminating boundaries in the space of original variables.

Out-of-fold cross-validation results confirmed the stability of the model (Accuracy= 0.93; Macro-F1= 0.94; Kappa = 0.90), showing consistency between training and generalisation. The small difference between cross-validation performance and independent test set performance indicates the absence of significant overfitting.

The analysis of variable importance revealed that the use of digital platforms (UP_DI), frequency of digital consumption (F_U), institutional perception (ER_UTN) and content assessment (SC_UTV) are the main factors distinguishing between profiles. These variables coincide with the structural dimensions identified in the MCA, reinforcing the internal consistency of the model.

The representative tree of the assembled model showed variables related to institutional consumption frequency (CF_U) and perception of UTV (SC_UTV) appearing in early division nodes, evidencing their structural role in cluster differentiation. Taken together, these results confirm that segmentation is not only statistically stable but also structurally interpretable.

5. Discussion

5.1. Digitisation and Intensity of Institutional Linkages

The results show that the differentiation between information profiles does not respond solely to digitisation patterns, but rather to the interaction between the use of digital platforms and the degree of institutional engagement. In this sense, variables such as the use of digital media (UP_DI) and the frequency of digital consumption (F_U) emerge as key structural factors in shaping information behaviour, which supports the findings of hypothesis H1. Likewise, the findings suggest that digitisation can coexist with different levels of institutional media valuation and consumption, giving rise to hybrid information profiles. This challenges the dichotomous configurations between digital and institutional media, and provides empirical evidence in favour of hypothesis H2.

 doi:10.20944/preprints202603.1050.v1

This finding is relevant in the context of university communication sustainability, as it challenges the traditional narrative that pits digital media against academic institutions. Rather than a linear shift towards digital, the data shows a recomposition of the information link, where the intensity of institutional commitment depends both on consumption experience and on the perception of quality and trust in content.

5.2. Theoretical Implications: Hybrid Digitisation and Institutional Sustainability

The results provide empirical evidence for the contemporary discussion on digital transformation in educational organisations, particularly in contexts where institutional sustainability increasingly depends on strategic communication management. Recent studies have pointed out that digitisation should not be interpreted as a process of substitution between traditional and digital channels, but rather as an ecosystemic reconfiguration of the information environment [70]. In line with this perspective, the findings show that the intensive use of digital platforms can coexist with different levels of institutional valuation, configuring hybrid consumption profiles that challenge the dichotomous logic of 'digital vs. institutional'.

The evidence supports approaches that view organisational communication as a central component of sustainability and institutional legitimacy [90]. The differentiation of profiles based on consumption frequency, institutional perception, and content evaluation suggests that communicational sustainability does not depend exclusively on technological adoption, but rather on the building of trust and symbolic capital in complex digital environments [91].

Beyond access to technological infrastructure, the differentiation between information profiles can be explained by usage patterns, levels of media exposure, and forms of interaction with institutional content, contributing to an understanding of the multidimensional nature of the digital divide and supporting hypothesis H5. This refers to the fact that the digital divide should be interpreted as a structural inequality within the information ecosystem, where differences in media literacy, consumption habits, and institutional trust directly influence individuals' ability to participate in contemporary information processes.

5.3. Hybrid Configurations and Tensions in Institutional Consumption

Structural analysis of the assembled model suggests that the differentiation between profiles does not respond to isolated variables, but rather to coherent combinations of digitisation and institutional valuation. This configurational logic aligns with recent approaches that conceptualise digital transformation as a systemic and multidimensional process, where technology, organisational culture and perception of value interact dynamically [70]. High digital intensity does not necessarily imply institutional displacement, but rather a reconfiguration of the informational bond, which coincides with studies that highlight the hybrid nature of contemporary media ecosystems [92].

Empirical evidence shows that individuals do not exclusively engage with traditional or digital media environments, but rather combine different information channels according to their consumption habits, interests, and levels of institutional trust. This behaviour reinforces hypothesis H4, which posits that information profiles within local media ecosystems are characterised by hybrid consumption configurations. In this sense, the local media ecosystem is configured as a dynamic space in which traditional media, digital platforms and institutional media converge, generating complex patterns of interaction that cannot be explained solely by technological substitution.

5.4. Structural Hierarchy of the Assembled Model

The aggregate analysis of the Random Forest model (Table 5) confirms the existence of a hierarchical architecture in profile differentiation. Permutation importance metrics identify the use of digital platforms (UP_DI) and the frequency of digital consumption (F_U) as the main discriminating factors. However, variables associated with institutional assessment, such as perception of UTV (SC_UTV), institutional evaluation (ER_UTN), and frequency of institutional consumption (CF_U) appear immediately after in predictive relevance.

Table 5. *Structural hierarchy of variables in the assembled model.*

Hierarchical level	Dominant variables	Empirical evidence
Level 1: Digitisation	UP_DI (use of digital platforms) F_U (Frequency of digital media use)	Greater importance by permutation (0.22 and 0.15)
Level 2: Institutional Linkage	SC_UTV (Tuning in to or consumption of university channels) ER_UTN (Exhibition at UTN university radio station) CF_U (Knowledge of institutional university media)	High MDI importance and early appearance of nodes
Level 3: Contextual refinement	CF_E (Consumption of printed or external media) CUTV (Specific consumption of UTV programmes) C_F1 (Preferred channel or primary format) QS_ML (Level of satisfaction with local media)	High frequency in deep splits

The convergence between permutation importance and impurity reduction (MDI), as well as frequency of use in forest nodes, reveals a tiered structure of differentiation. At the first level, the model separates profiles according to digital intensity; at the second level, it segments according to institutional linkages; and at subsequent levels, it refines the classification using specific content variables. This hierarchy confirms that digitisation acts as the primary axis, but the consolidation of profiles depends on the interaction between digital behaviour and institutional capital.

The early appearance of variables associated with institutional perception and university media consumption within the model structure reveals that greater exposure to these media is linked to more favourable evaluations of academic management and communication content. This provides evidence for hypothesis H3. This finding emphasises that the construction of institutional trust is closely related to the frequency of media interaction, consolidating the role of university media as relevant actors within the territorial information ecosystem.

Finally, the high performance of the machine learning model provides evidence for hypothesis H6, demonstrating that variables related to digitisation and institutional consumption have a high predictive capacity for classifying differentiated information profiles. The consistency between unsupervised segmentation and the results of the supervised model suggests that the patterns identified respond to real structures of information behaviour within the media ecosystem analysed. This result complements the potential of machine learning as an analytical tool for understanding complex communication dynamics in information sustainability studies.

6. Limitations and future lines of research

The analysis is limited to a specific geographical context, the province of Imbabura in Ecuador, which limits the possibility of generalising the findings to other territorial contexts with different socio-economic, cultural or media characteristics. Similarly, the research is based on cross-sectional data, which restricts the ability to observe changes over time in media consumption patterns and in the relationship between digitisation and institutional engagement.

Although the sample size allows for the identification of robust segmentation patterns, the information is based on participants' self-reported perceptions, which may be influenced by memory bias or social desirability bias. Furthermore, the analytical model focuses primarily on variables related to media consumption and institutional perception, so other potentially relevant factors, such as digital skills, cultural capital, or levels of media literacy, were not explicitly incorporated into the analysis.

In methodological terms, although the combined use of multivariate analysis and machine learning techniques made it possible to identify complex structures in information behaviour, the study focused on a specific set of algorithms and variables, opening up the possibility of exploring additional analytical approaches that delve deeper into the interpretation of these phenomena. In this regard, future research could incorporate more advanced machine learning models or explanatory techniques that allow for a more detailed analysis of the causal relationships between digitisation, institutional trust and information sustainability.

Finally, future lines of research could broaden the scope of the study through longitudinal designs that allow for the analysis of the evolution of information profiles over time, as well as interregional or international comparisons that contribute to understanding how media ecosystems vary in different territorial contexts.

7. Conclusions

The results confirm that information sustainability in local media ecosystems does not depend solely on access to digital technologies, but also on how individuals integrate into the information system through different patterns of media consumption and levels of institutional engagement. The identification of differentiated audience profiles shows that the digital divide is a multidimensional phenomenon that transcends the availability of technological infrastructure, as it also involves usage practices, levels of information exposure, and degrees of trust in institutions. In this sense, the findings show that digitisation acts as a structuring axis of information behaviour, but that the consolidation of profiles depends on the interaction between media consumption, institutional valuation and forms of participation within the communication ecosystem.

From a theoretical perspective, the research contributes to the field of social sustainability by proposing information sustainability as a structural dimension of territorial development within local media ecosystems. This approach broadens our understanding of the digital divide by integrating variables related to media consumption, institutional perception, and the interaction between traditional, digital, and institutional media. Likewise, the study contributes to the literature on media ecosystems by demonstrating that information patterns are shaped by hybrid consumption dynamics, in which digitisation does not completely replace traditional or institutional media, but rather articulates with them within an interdependent information system. This perspective is particularly relevant for Latin American contexts, where studies on information sustainability and audience segmentation are still limited.

In methodological terms, the work demonstrates the potential of integrating multivariate segmentation techniques with machine learning models to analyse complex communication phenomena. The combination of multiple correspondence analysis, hierarchical classification, and Random Forest models made it possible to identify latent structures in information behaviour and validate the stability of the profiles found. This methodological approach expands the tools available for studying institutional communication and information sustainability by offering greater explanatory power to analyse non-linear interactions between variables related to digitalisation, media consumption, and institutional trust.

From a practical and public policy perspective, the results suggest that strategies aimed at reducing the digital divide must go beyond the expansion of technological infrastructure. In particular, universities, local governments and regional media should implement media and digital literacy policies that strengthen audiences' critical capacities to interact with contemporary information ecosystems. Likewise, university media can play a strategic role in promoting information sustainability by developing accessible content, multi-platform formats, and segmented communication strategies that respond to the different consumption profiles identified.

doi:10.20944/preprints202603.1050.v1

In addition, it is recommended that public and educational institutions promote institutional communication programmes based on data analysis that allow for a more accurate understanding of the information dynamics of their audiences. This includes the use of advanced analytical tools to monitor media consumption patterns, design differentiated territorial communication strategies, and strengthen information transparency. Similarly, strengthening partnerships between universities, local media, and community organisations could contribute to creating more inclusive information environments, promoting citizen participation and equitable access to reliable information

Overall, this study demonstrates that information sustainability is a key component of social sustainability in territories, as it directly influences the way people access, interpret and use information in their daily lives. Understanding the different configurations of media consumption allows for the design of more inclusive and effective communication strategies aimed at strengthening institutional trust and reducing information inequalities. In this sense, the integration of analytical approaches based on machine learning opens up new opportunities for the development of evidence-based public policies and institutional strategies capable of promoting more sustainable, participatory and socially integrated media ecosystems.

Author Contributions: Conceptualization, L.S.-E. and J.P.-P.; methodology, D.E.C.-T; L.S.-E. and J.P.-P; software, D.E.C.-T.; validation, D.E.C.-T. and L.S.-E.; formal analysis, D.E.C.-T. and L.S.-E.; investigation, L.S.-E.; J.P.-P. and D.E.C.-T.; resources, L.S.-E. and D.E.C.-T.; data curation, D.E.C.-T. and L.S.-E; writing—original draft preparation, L.S.-E.; J.P.-P. and D.E.C.-T.; writing—review and editing, L.S.-E. and J.P.-P.; visualization, D.E.C.-T; supervision, L.S.-E. and J.P.-P; project administration, L.S.-E. and J.P.-P. All authors have read and agreed to the published version of the manuscript.

Funding: The APC was financed by the Universidad Técnica del Norte. Code: InvestigaUTN-2025-1499.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: All subjects who participated in the study were informed about the study and gave their express consent.

Data Availability Statement: The research instrument for this study is available at https://doi.org/10.17605/OSF.IO/JDK7U.

Acknowledgments: The authors express their gratitude for the support received from Universidad Técnica del Norte.

Conflicts of Interest: The authors state that they have no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Muthu, S.S. Assessment of Social Sustainability Management in Various Industrial Sectors. In *Environmental Footprints and Eco-Design of Products and Processes*; Springer: Cham, Switzerland, **2025**; pp. 1–4. https://doi.org/10.1007/978-3-031-81029-9_1
2. Beyer, K.; Arnold, M.G. Examining the social side of sustainability in the debate on business model innovations in the textile, clothing and fashion industry: A typology based on the value chain perspective. *Int. J. Innov. Sustain. Dev.* **2022**, 16, 322–371. https://doi.org/10.1504/IJISD.2022.123908
3. Segovia-Martin, J.; Tamariz, M. Synchronising institutions and value systems: A model of opinion dynamics mediated by proportional representation. *PLoS ONE* **2021**, 16, e0257525. https://doi.org/10.1371/journal.pone.0257525
4. Papadopoulos, T.; Broadbent, R. Reducing the digital divide: An evaluation framework for the wired community@collingwood project. *Proc. Int. Conf. Inf. Soc*. **2009**, pp. 701–709.

 doi:10.20944/preprints202603.1050.v1

5. Etim, A.S. Adoption and Use of Technology Tools and Services by Economically Disadvantaged Communities: Implications for Growth and Sustainability. In *Digital Inclusion and Economic Development*; IGI Global: Hershey, PA, USA, **2023**; p. 383. https://doi.org/10.4018/978-1-6684-5347-6
6. Cerrillo Vidal, J.A.; Beluschi-Fabeni, G. Resistances, difficulties and fears: An approach to understanding the causes of persistent digital divides. *Eur. Public Soc. Innov. Rev.* **2024**, 9. https://doi.org/10.31637/epsir-2024-603
7. García Zare, E.J.; Soto Abanto, S.E.; Rodríguez Paredes, N.P.; Merino Salazar, T.D.R.; Pagador Flores, S.E.; Baldárrago Baldárrago, J.L.; Salas-Ruiz, J.A.; Mejía Pardo, P.I. Technological Devices and Digital Competences: A Look into the Digital Divides for University Continuity during the COVID-19 Pandemic. *Sustainability* **2023**, 15, 18494. https://doi.org/10.3390/su15118494
8. Khlouf, M.M.M.; Azmoty, M.N. From FM to TikTok: A mixed-methods study of platform gaps, content misalignment, and rural–urban divides in Jordanian university radio engagement. *Front. Commun.* **2025**, 10. https://doi.org/10.3389/fcomm.2025.1678387
9. Suing, A.; Suárez-Villegas, J. Perfil, incidencia y perspectivas de la desinformación entre los ecuatorianos. *Journal. Media* **2024**. https://doi.org/10.3390/journalmedia5030063
10. Villamar, I.; Granda, L.; Viejó, J.; Cobos, J. Impact of information behavior on marketing and fertilization strategies of small cocoa producers in Ecuador. *Agriculture* **2025**. https://doi.org/10.3390/agriculture15080858
11. Luzuriaga-Amador, M.; Novillo-Luzuriaga, N.; Guevara-Viejó, F.; Valenzuela-Cobos, J. Evaluation of information competencies in fertilization and commercialization strategies of small banana producers in Ecuador. *Sustainability* **2025**, 17, 30868. https://doi.org/10.3390/su17030868
12. Caviedes, Y.; López, L.; Prieto, A. Las redes sociales como escenario en el que se construyen los imaginarios políticos de los jóvenes: Caso Ecuador. *Signo Pensam.* **2025**. https://doi.org/10.11144/javeriana.syp44.rsec
13. Capriotti, P.; Losada-Díaz, J.; Martínez-Gras, R. Evaluating the content strategy developed by universities on social media. *Prof. Inf.* **2023**. https://doi.org/10.3145/epi.2023.mar.10
14. Capriotti, P.; Zeler, I. Análisis de la comunicación efectiva en redes sociales en instituciones de educación superior. *Humanit. Soc. Sci. Commun.* **2023**, 10. https://doi.org/10.1057/s41599-023-02187-8
15. Galioto, M.; Pedone, F.; Vantarakis, A.; Tavares, P.; Bianco, A. The use of social media in university institutional communication: A prospective model to strengthen students' sense of belonging. *Front. Commun.* **2025.** https://doi.org/10.3389/fcomm.2025.1523295
16. Castillo, L.; Alarcón-Llontop, L.R. Plan de comunicación basado en tecnologías de la información y la comunicación para la cultura organizacional universitaria peruana. *Edelweiss Appl. Sci. Technol.* **2025**. https://doi.org/10.55214/25768484.v9i5.7586
17. Nishan, A.; Hossain, M.; Haque, M.; Ankhi, R.; Hosen, A. The intersection of macroeconomic narratives and public trust. *Cognizance J. Multidiscip. Stud.* **2025**. https://doi.org/10.47760/cognizance.2025.v05i04.002
18. Guo, S.; Wang, D.; Shen, C. The gap between expectations and evaluation: Audience interaction with news and perceptions of media performance and credibility. *Glob. Media China* **2024**. https://doi.org/10.1177/20594364241268137
19. Al-Mshashaqbeh, Y. Modeling Jordanian audience engagement and news credibility in the digital media era: A PLS-SEM approach. *Front. Commun.* **2025**. https://doi.org/10.3389/fcomm.2025.1675693
20. Choi, Y.; Lee, C. Profiling AI speaker users: Machine learning insights into consumer adoption patterns. *PLoS ONE* **2024**, 19. https://doi.org/10.1371/journal.pone.0315540
21. Zeybek, S.; Kaçaman, M. User interest classification on social networks using machine learning and deep learning. *Ömer Halisdemir Univ. J. Eng. Sci.* **2025**. https://doi.org/10.28948/ngumuh.1655764
22. Ali, A.; Raza, A.; Sayed, M.; Qureshi, B.; Memon, Y. Data-driven machine learning approaches for Netflix content analysis and visualization. *J. Eng. Res. Rep.* **2025**, 27. https://doi.org/10.9734/jerr/2025/v27i41471
23. Fayq, K.; Tkatek, S.; Idouglid, L. Improving TV program success prediction through machine learning by integrating people-meter and digital interaction metrics. *Indones. J. Electr. Eng. Comput. Sci.* **2025**. https://doi.org/10.11591/ijeecs.v39.i1.pp353-363
24. Cruz, J. Modelos predictivos basados en aprendizaje automático para el análisis del comportamiento digital. *Datos Metadatos* **2025**. https://doi.org/10.56294/dm2025994

25. Hall, T.D. Information Redlining: The urgency to close the digital access and literacy divide and the role of libraries as lead interveners. *J. Libr. Adm.* **2021**, 61, 484–492. https://doi.org/10.1080/01930826.2021.1906559
26. Rydzewski, P. Digital inequality and sustainable development. *Probl. Ekorozw.* **2025**, 20, 96–108. https://doi.org/10.35784/preko.6691
27. Aalberg, T.; Curran, J. How media inform democracy: Central debates. In *How Media Inform Democracy: A Comparative Approach*; Routledge: London, UK, **2012**; pp. 3–14
28. Elenbaas, M.; de Vreese, C.; Schuck, A.; Boomgaarden, H. Reconciling passive and motivated learning: The saturation-conditional impact of media coverage and motivation on political information. *Commun. Res.* **2014**, 41, 481–504
29. Perry, H.B. Is access enough? Interrogating the influence of money and power in shaping information. *Open Inf. Sci.* **2020**, 4, 29–38
30. Erek, K.; Schmidt, N.H.; Zarnekow, R.; Kolbe, L.M. Management instruments for sustainable information systems management. In *Green Technologies: Concepts, Methodologies, Tools and Applications*; IGI Global: Hershey, PA, USA, **2011**; pp. 1448–1465
31. Ștefana-Cătălina, P.-D.; Ionuț-Viorel, H.; Marius, P.; Petronela, L.R. Enhanced model for evaluating the information systems success from the sustainability management perspective. In *Green Intelligence: Carbon Footprint Assessment in Digital Economy*; CRC Press: Boca Raton, FL, USA, **2025**; pp. 57–73. https://doi.org/10.1201/9781003518990-5
32. Erek, K. From green IT to sustainable information systems management: Managing and measuring sustainability in IT organizations. In *Proceedings of the Americas Conference on Information Systems (AMCIS)*, Detroit, MI, USA, **2011**; pp. 766–781
33. Selvakumar, P. The role of technology in enhancing social sustainability. In *Enhancing Social Sustainability in Manufacturing Supply Chains*; IGI Global: Hershey, PA, USA, **2025**; pp. 275–302. https://doi.org/10.4018/979-8-3693-9740-4.ch010
34. Unwin, T.; de Bastion, G. Digital divide. In *International Encyclopedia of Human Geography*; Elsevier: Oxford, UK, **2009**; pp. 191–197. https://doi.org/10.1016/B978-008044910-4.00090-0
35. Hilbert, M. Digital divide(s). In *The International Encyclopedia of Digital Communication and Society*; Wiley: Hoboken, NJ, USA, 2015; pp. 1–7. https://doi.org/10.1002/9781118767771.wbiedcs012
36. Huang, T.; Quan, Y. Narrowing the digital divide: The growth and distributional effect of internet use on income in rural China. *China Econ. Rev.* **2025**, 91, 102387. https://doi.org/10.1016/j.chieco.2025.102387
37. Hwang, Y.; Park, N. Digital divide in social networking sites. *Int. J. Mob. Commun.* **2013,** 11, 446–464. https://doi.org/10.1504/IJMC.2013.056955
38. Friemel, T.N. The digital divide has grown old: Determinants of a digital divide among seniors. *New Media Soc.* **2016**, 18, 313–331. https://doi.org/10.1177/1461444814538648
39. Huang, M.; Ye, Y. “A matter of life and death”: Mitigating the gray digital divide in using health information technologies in the post-pandemic era. *Health Commun.* **2025**, 40, 620–630. https://doi.org/10.1080/10410236.2024.2358279
40. Fuchs, C.; Horak, E. Africa and the digital divide. *Telemat. Inform.* **2008**, 25, 99–116. https://doi.org/10.1016/j.tele.2006.06.004
41. Ruohonen, J.; Tuikka, A.-M. Digital divides and online media. In *Proceedings of the ACM Conference on Information Technology for Social Good*, Roma, Italy, **2021**; pp. 157–163. https://doi.org/10.1145/3485768.3485815
42. Dubey, P.; Dubey, P.; Sahu, K.K. Bridging the gap: Unveiling the impact of COVID-19 and social media on the digital divide. In *Psychology and Intricacies in Social Media Interactions: A Critical Perspective*; Routledge: London, UK, **2025**; pp. 193–206. https://doi.org/10.1201/9781003614326-11
43. Otokiti, A.; Williams, K.S.; Warsame, L. Impact of digital divide on the adoption of online patient portals for self-motivated patients. *Healthc. Inform. Res.* **2020**, 26, 220–228. https://doi.org/10.4258/hir.2020.26.3.220
44. Wei, L.; Hindman, D.B. Does the digital divide matter more? Comparing the effects of new media and old media use on the education-based knowledge gap. *Mass Commun. Soc.* **2011**, 14, 216–235. https://doi.org/10.1080/15205431003642707

45. Abdelfattah, B.M. Individual-multinational study of internet use: The digital divide explained by displacement hypothesis and knowledge-gap hypothesis. *World Appl. Sci. J.* **2012**, 5, 3649–3657
46. Netto, V.M.; Vargas, J.C.; Saboya, R.T. The social effects of architecture: Built form and social sustainability. In *Urban Social Sustainability: Theory, Policy and Practice*; Routledge: London, UK, **2019**; pp. 125–148. https://doi.org/10.4324/9781315115740
47. Brown, M.; Marmarà, V. 'Media-ted' electoral campaigns: Europeanisation and postcolonial dynamics of voters' use of media platforms in Malta. *Postcolonial Dir. Educ.* **2022**, 11, 42–78
48. Weston, M.; Vogler, D.; Jürgens, P.; Rauchfleisch, A.; Eisenegger, M. Tracking the roots of low news usage on smartphones: Do individual interests, political attitudes, and media consumption habits affect the mobile news use of young adults? *J. Broadcast. Electron. Media* **2025**, 69, 529–548. https://doi.org/10.1080/08838151.2025.2535389
49. Wang, H.; Cai, T. Media exposure and Chinese college students' attitudes toward China's maritime claims and disputes in the South and East China Seas. *Cogent Soc. Sci.* **2018**, 4, 1482995. https://doi.org/10.1080/23311886.2018.1482995
50. Li, T.; Zhu, H. Effect of the media on the opinion dynamics in online social networks. *Physica A* **2020**, 551, 124117. https://doi.org/10.1016/j.physa.2019.124117
51. Out, C.; Tu, S.; Neumann, S.; Zehmakan, A.N. The impact of external sources on the Friedkin-Johnsen model. In *Proceedings of the ACM Web Conference*, Singapore, **2024**; pp. 1815–1824. https://doi.org/10.1145/3627673.3679780
52. Stehr, P.; Rössler, P.; Leissner, L.; Schönhardt, F. Parasocial opinion leadership: Media personalities' influence within parasocial relations. *Int. J. Commun.* **2015**, 9, 982–1001
53. Asher, D.E.; Caylor, J.P.; Neigel, A.R. Effects of social media involvement, context, and data-type on opinion formation. In *Proceedings of the IEEE International Conference on Social Computing*, **2018**; pp. 32–37. https://doi.org/10.1109/SocialSens.2018.00019
54. Johnson, T.; Kellstedt, P.M. Media consumption and the dynamics of policy mood. *Polit. Behav.* **2014**, 36, 377–399. https://doi.org/10.1007/s11109-013-9233-5
55. Mulyani, M.; Octavia, J.R. Educational board game design based on local wisdom for youth to support social sustainability learning. *IOP Conf. Ser. Earth Environ. Sci.* **2023**, 1256, 012006. https://doi.org/10.1088/1755-1315/1256/1/012006
56. Wolff, L.-A.; Ehrström, P. Social sustainability and transformation in higher educational settings: A utopia or possibility? *Sustainability* **2020**, 12, 4176. https://doi.org/10.3390/su12104176
57. Ispir, B. Interactive media steer in educational television programs. In *Digital Multimedia: Concepts, Methodologies, Tools, and Applications*; IGI Global: Hershey, PA, USA, **2017**; pp. 1111–1122. https://doi.org/10.4018/978-1-5225-3822-6.ch053
58. Gori, E.; Romolini, A.; Fissi, S.; Contri, M. Toward the dissemination of sustainability issues through social media in the higher education sector: Evidence from an Italian case. *Sustainability* **2020**, 12, 4658. https://doi.org/10.3390/su12114658
59. Atakan, S.; İşcioğlu, T.E. Sustainability practices of higher education institutions—An analysis from a developing country. In *Accounting, Finance, Sustainability, Governance and Fraud*; Springer: Singapore, **2019**; pp. 215–235. https://doi.org/10.1007/978-981-13-7924-6_12
60. Sart, G. Sustainable campus design in universities. In *Considerations on Education for Economic, Social, and Environmental Sustainability*; IGI Global: Hershey, PA, USA, **2023**; pp. 121–135. https://doi.org/10.4018/978-1-6684-8356-5.ch006
61. Nicolino, G.; Barros, S. Introducing the graphical assessment of universities' sustainability image (GAUSI) instrument: A marketing tool. In *World Sustainability Series*; Springer: Cham, Switzerland, **2016**; pp. 193–212. https://doi.org/10.1007/978-3-319-26866-8_12
62. Carretón-Ballester, C.; Quiles-Soler, C.; Lorenzo-Solá, F. Social responsibility of Spanish universities for sustainable relationships. *Prof. Inf.* **2023**, 32. https://doi.org/10.3145/epi.2023.nov.02
63. Shriberg, M. Is the "maize-and-blue" turning green? Sustainability at the University of Michigan. *Int. J. Sustain. High. Educ.* **2003**, 4, 263–276. https://doi.org/10.1108/14676370310485465

 doi:10.20944/preprints202603.1050.v1

64. Unz, D.C. Exposure to news. In *The International Encyclopedia of Communication*; Wiley: Malden, MA, USA, **2008**. https://doi.org/10.1002/9781405186407.wbiece060
65. Sobkowicz, P.; Kaschesky, M.; Bouchard, G. Opinion mining in social media: Modeling, simulating, and forecasting political opinions in the web. *Gov. Inf. Q.* **2012**, 29, 470–479. https://doi.org/10.1016/j.giq.2012.06.005
66. McLuhan, M. *Understanding Media: The Extensions of Man*; McGraw-Hill: New York, NY, USA, **1964**.
67. Postman, N. *Teaching as a Conserving Activity*; Delta: New York, NY, USA, **1979**
68. Napoli, P. Exposure diversity reconsidered. *J. Inf. Policy* **2011**. https://doi.org/10.5325/jinfopoli.1.2011.0246
69. Nambisan, S.; Zahra, S.A.; Luo, Y. Global platforms and ecosystems: Implications for international business theories. *J. Int. Bus. Stud.* **2019**, 50, 1464–1486. https://doi.org/10.1057/s41267-019-00262-4
70. Verhoef, P.C.; Broekhuizen, T.; Bart, Y.; Bhattacharya, A.; Dong, J.Q.; Fabian, N.; Haenlein, M. Digital transformation: A multidisciplinary reflection and research agenda. *J. Bus. Res.* **2021**, 122, 889–901. https://doi.org/10.1016/j.jbusres.2019.09.022
71. Atkinson, W. Charting fields and spaces quantitatively: From multiple correspondence analysis to categorical principal components analysis. *Qual. Quant.* **2024**, 58, 829–848. https://doi.org/10.1007/s11135-023-01669-w
72. Matta, E.; Stalidis, G. Profiling online and physical supermarket customers using factor and clustering methods. *Smart Innov. Syst. Technol.* **2024**, 386, 227–243. https://doi.org/10.1007/978-981-97-1552-7_15
73. Saráuz-Estevez, L.; Pupiales-Proaño, J.; Cuaical-Tapia, D. Key factors in the sustainable growth of MSMEs in Ibero-America: An empirical study based on machine learning. *Sustainability* **2026**, 18, 1940. https://doi.org/10.3390/su18041940
74. D'Enza, A.I.; Markos, A. Incremental visualization of categorical data. In *Advances in Statistical Models for Data Analysis*; Morlini, I., Minerva, T., Vichi, M., Eds.; Springer: Cham, Switzerland, **2015**; pp. 137–148. https://doi.org/10.1007/978-3-319-17377-1_15
75. Khangar, N.V.; Kamalja, K.K. Multiple correspondence analysis and its applications. *Electron. J. Appl. Stat. Anal.* **2017**, 10, 432–462. https://doi.org/10.1285/i20705948v10n2p432
76. Agapito, D.; Valle, P.; Mendes, J. The sensory dimension of tourist experiences: Capturing meaningful sensory-informed themes in Southwest Portugal. *Tour. Manag.* **2014**, 42, 224–237. https://doi.org/10.1016/j.tourman.2013.11.011
77. Greenacre, M.; Groenen, P.; Hastie, T.; Iodice D'Enza, A.; Markos, A.; Tuzhilina, E. Principal component analysis. *Nat. Rev. Methods Primers* **2022,** 2, 100. https://doi.org/10.1038/s43586-022-00184-w
78. Le Roux, B.; Rouanet, H. *Multiple Correspondence Analysis*; SAGE Publications: Thousand Oaks, CA, USA, **2010**. https://doi.org/10.4135/9781412993906
79. Costa, P.S.; Santos, N.C.; Cunha, P.; Cotter, J.; Sousa, N. The use of multiple correspondence analysis to explore associations between categories of qualitative variables in healthy ageing. *J. Aging Res.* 2013, **2013**, 302163. https://doi.org/10.1155/2013/302163
80. Beh, E.; Lombardo, R. Correspondence analysis using the Cressie–Read family of divergence statistics. *Int. Stat. Rev.* **2024**, 92, 17–42. https://doi.org/10.1111/insr.12541
81. Murtagh, F.; Legendre, P. Ward's hierarchical agglomerative clustering method: Which algorithms implement Ward's criterion? *J. Classif.* **2014,** 31, 274–295. https://doi.org/10.1007/s00357-014-9161-z
82. Rani, P.; Sid Ahmed, N.; Jain, A. A decision support system for heart disease prediction based upon machine learning. *J. Reliab. Intell. Environ.* **2021**, 7. https://doi.org/10.1007/s40860-021-00133-6
83. Xu, D.; Tian, Y. A comprehensive survey of clustering algorithms. *Ann. Data Sci.* **2015,** 2, 165–193. https://doi.org/10.1007/s40745-015-0040-1
84. Breiman, L. Random forests. *Mach. Learn.* 2001, 45, 5–32. https://doi.org/10.1023/A:1010933404324
85. Probst, P.; Boulesteix, A.-L.; Wright, M.N. Hyperparameters and tuning strategies for random forest. *arXiv* **2018**, arXiv:1804.03515. https://doi.org/10.48550/arXiv.1804.03515
86. Lundberg, S.M.; Erion, G.; Chen, H.; DeGrave, A.; Prutkin, J.; Nair, B.; Katz, R.; Himmelfarb, J.; Bansal, N.; Lee, S.-I. From local explanations to global understanding with explainable AI for trees. *Nat. Mach. Intell.* **2020,** 2, 56–67. https://doi.org/10.1038/s42256-019-0138-9

 doi:10.20944/preprints202603.1050.v1

87. Biau, G.; Scornet, E. A random forest guided tour. *TEST* **2016,** 25, 197–227. https://doi.org/10.1007/s11749-016-0481-7
88. Kumar, S.; Kaur, P.; Gosain, A. A comprehensive survey on ensemble methods. In *Proceedings of the International Conference on Innovative Trends in Communication and Computer Technology (I2CT)*, **2022.** https://doi.org/10.1109/I2CT54291.2022.9825269
89. Chicco, D.; Jurman, G. The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC Genomics* **2020**, 21, 6. https://doi.org/10.1186/s12864-019-6413-7
90. Capriotti, P.; Zeler, I.; Camilleri, M.A. Corporate communication through social networks: The identification of the key dimensions for dialogic communication. In *Strategic Corporate Communication in the Digital Age*; Emerald Publishing: Bingley, UK, **2021**; pp. 33–52. https://doi.org/10.1108/978-1-80071-264-520211003
91. Salvi, A.; Vitolla, F.; Rubino, M.; Giakoumelou, A.; Raimo, N. Online information on digitalisation processes and its impact on firm value. *J. Bus. Res.* **2021**, 124, 437–444. https://doi.org/10.1016/j.jbusres.2020.10.025
92. Nambisan, S.; Wright, M.; Feldman, M. The digital transformation of innovation and entrepreneurship: Progress, challenges and key themes. *Res. Policy* **2019,** 48, 103773. https://doi.org/10.1016/j.respol.2019.03.018