

# A Data-Driven Method for Cross-Regional User Insight

Shiqi Yao<sup>1</sup> and Danhua Zhao<sup>2</sup>

<sup>1</sup>School of design, Hunan University, Changsha 410082, China

<sup>2</sup>Research Institute of HNU in Chongqing, Chongqing, 401135, China

## ABSTRACT

In the context of globalization, products and services are increasingly designed for users from diverse cultural backgrounds. How to generate user insights that are both interpretable and translatable across different social structures, cultural psychologies, and institutional environments has become a central methodological challenge in cross-regional user research. Existing studies often exhibit a disconnection between macro-level national or contextual analysis and micro-level user experience insights: macro analyses are difficult to translate into concrete design decisions, while experience-based data analysis frequently lacks systematic interpretation within institutional and cultural contexts. To address this gap, this paper proposes a data-driven method for cross-regional user insight based on the integration of multi-source data and a layered analytical framework. The proposed method consists of three interrelated layers. At the practical layer, macro statistical data and social event frequency analysis are used to quantitatively examine material environments and survival constraints across regions. At the theoretical layer, natural language processing (NLP) techniques are applied to user-generated content to analyze emotional tendencies and value semantics, revealing culturally embedded psychological perceptions. At the strategic layer, the AEIOU framework is employed to structurally analyze everyday practices and infer broader lifestyle paradigms and institutional influences from the bottom up. Triangulation is introduced as the core analytical strategy to independently encode and cross-validate findings across the three layers, enabling the identification of cross-regionally consistent core needs as well as potential design opportunities exposed by structural differences. The concept of “comfort” is used as an application context to demonstrate the feasibility of the proposed method. The results indicate that this approach enhances both the explanatory depth and application stability of cross-regional user insights, providing methodological support for user research and decision translation in global design practices.

**Keywords:** Cross-regional user insight, Big data user research, Cross-cultural research methods, Sentiment analysis, Triangulation-based validation

## INTRODUCTION

With the globalization of the digital economy, the user base of products and services has exhibited cross-regional cultural diversity. Users’ cultural backgrounds profoundly influence their product experiences and behavioral preferences: users from different cultural backgrounds have different focuses and needs for the same product. Therefore, cross-cultural user research is

becoming increasingly important. However, traditional user research is often limited to local small-sample investigations, lacks a cross-cultural perspective, and makes it difficult to comprehensively grasp the diverse needs of users in other regions (Lee & Park, 2023).

The rise of big data provides new opportunities for cross-cultural user research. Massive user-generated content and behavioral data enable researchers to conduct large-scale quantitative analysis of global user behavior (Wu et al., 2023). Methods that combine big data text mining with cross-cultural questionnaires have been used to characterize cognitive differences among users in different markets, and results show that this approach can provide more comprehensive user insights. Traditional interviews and questionnaires are constrained by subjective bias and sample limitations, while a single big-data analysis is difficult to capture details of user behavior; multi-source data collaboration is expected to balance breadth and depth, compensate for the shortcomings of a single method, and enhance the reliability of conclusions (Wu et al., 2025). However, at present, user research methods that integrate big data in cross-cultural contexts are still imperfect: existing studies either focus on comparative analysis of only a few cultures, or lack a unified framework for explaining complex contexts (Lee, 2025). This paper attempts to take the integration of a cross-cultural perspective with big data and multi-source data analysis as an entry point, proposing an operational and adaptable user research pathway to support user insights in complex cultural contexts.

## LITERATURE REVIEW

Currently, cross-cultural user research targeting different countries in academia and industry mainly follows two paths.

The first path is anthropology-based qualitative research, whose core methods include field ethnography, participatory observation, and in-depth interviews. The advantage of this method lies in its ability to obtain “thick data” and deeply understand the cultural context behind user behavior (Wang, 2016). However, its limitation is “micro-level myopia”: researchers often immerse themselves in individual narratives while ignoring the massive institutional forces that shape these behaviors. In addition, traditional field research is extremely costly and is easily influenced by researchers’ own ethnocentrism (Karhapää et al., 2025).

Another user research path is data-science-based quantitative analysis, which leverages large-scale data such as user-generated content, social media records, and click logs to provide high-frequency, scalable, real-time insights for user research (Abkenar et al., 2021). Text mining and sentiment analysis are widely applied to unstructured texts such as user reviews to extract valuable information, thereby overcoming the scale limitations of traditional qualitative methods. However, this type of method also faces challenges: on the one hand, it is often difficult to capture cultural context and rhetorical usage—irony, idioms, and cultural details in online communication often exceed algorithmic understanding (Roy, 2025); on the other hand, these models typically rely on predefined label systems or prior classifications,

ignoring the ambiguity and context dependence of language, which may lead to semantic drift or algorithmic bias (Guo et al., 2024).

In recent years, some studies have begun to try combining multi-source data with cultural-psychological modeling to enhance the explanatory power and transferability of cross-cultural analysis. For example, (Wu et al., 2023) proposed introducing cultural psychology theory into a big data analysis framework, exploring the long-term influence mechanism of culture on individual behavior through the combination of micro-level social behavior and macro-level psychological structure. They pointed out that in cross-cultural research, “heterogeneity” and “coupling” coexist among data sources; it is necessary to build layered structures and cross-validation mechanisms to strengthen the explanatory connections among data.

Current multi-source data fusion methods are mostly concentrated on specific analytical strategies or improving the precision of a certain data type, and still lack a systematic cross-cultural user research pathway to integrate the linkage mechanisms among macro statistics, meso-level semantic structures, and micro-level subjective experiences. Therefore, on the basis of existing methodological experience, further constructing a user research scheme suitable for complex cultural structures and multi-source data collaboration still entails practical challenges and theoretical space.

## RESEARCH METHOD

### Construction of the Three-Layer Analytical Framework

To scientifically deconstruct user experience concepts characterized by high contextual dependency and cultural polysemy, this study draws on the ecological systems theory proposed by Bronfenbrenner (1979). This theory emphasizes that individual development and experience are influenced not only by the immediate interactive environment but also constrained by higher-level social structures and value systems, such as cultural norms and policy frameworks. It thus provides a theoretical basis for understanding how “macro-level strategic intentions penetrate and shape individual experiences.”

Based on this perspective, we define the research object as three interrelated and nested layers:

#### The Practical Layer

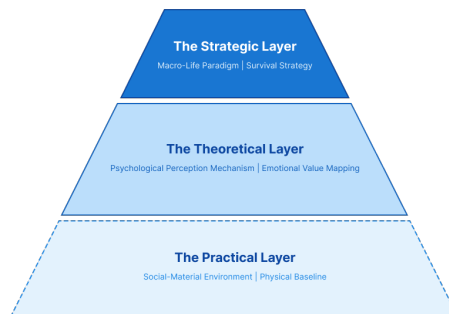
This layer represents the system’s “explicit manifestations.” According to Bourdieu’s (1977) theory of habitus, culture is first expressed through concrete and observable daily practices and material forms. As described by Reckwitz (2002) in social practice theory, practices are assemblages of actions, objects, and symbols. This corresponds to research elements such as “local consumption habits,” “architectural styles,” and “natural environments” that are observable in the field.

## The Theoretical Layer

This layer represents the system’s “interpretive mechanism.” According to Schein’s (2010) organizational culture model, beneath explicit behaviors and artifacts lie deeper layers of value judgments and basic assumptions, which determine how individuals perceive their environment, interpret experiences, and form emotional responses. This layer requires the introduction of psychological and sociocultural theories to reveal the emotional logic and value-semantic structures on which “comfort,” as a subjective experience, depends in different cultural contexts.

## The Strategic Layer

This layer represents the system’s “macro-level determining force.” According to Wallerstein’s (2004) world-systems analysis, micro-level phenomena within a region are ultimately constrained by its position and survival strategy within the macro political–economic structure.



**Figure 1:** The three-layer analytical framework.

Through these three layers, we construct a complete analytical loop of “practice (material reality) – theory (psychological demand) – strategy (macro constraint)” (see Figure 1).

Given that perceptions of “comfort” exhibit high ambiguity and structural variation across cultural contexts, this paper adopts comfort as the application scenario for the methodology. By analyzing comfort-related samples, the effectiveness of this framework in handling complex sensory indicators and uncovering deep cultural motivations is validated.

## ANALYTICAL FRAMEWORK AND PATHWAYS

### Bottom Layer (Practical Layer): Quantitative Scanning of the Social and Material Survival Environment

The practical layer focuses on the social and material survival environment of the target region. By analyzing macro statistical data and social events, this layer establishes a physical baseline for user experience.

Taking “comfort” perception as an example, the analysis concentrates on two dimensions: residential space constraints and economic consumption constraints.

Core indicators include per capita living area, population density, environmental noise, median disposable income, and the Engel coefficient. Kleeman & Foster (2025) found that housing area, layout flexibility, natural lighting, and ventilation are closely related to residential satisfaction under space-constrained conditions, indicating that spatial organization is a key physical factor influencing comfort experience. According to the OECD (2024), income-related indicators determine whether users can only afford survival-oriented functions or are capable of paying a premium for emotional experiences.

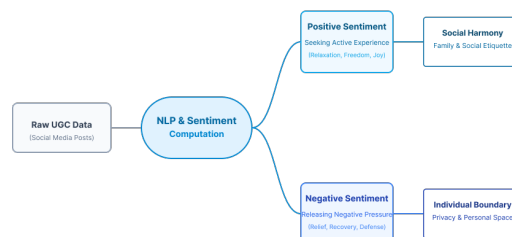
In terms of data acquisition, this study combines macro statistics with social event heat monitoring. While population and economic data are obtained from official statistical sources, dynamic social pressure is identified through the frequency of negative events. Hong (2025) demonstrated that high-frequency mentions of issues such as housing shortages, traffic congestion, and cost-of-living pressures in news reports can accurately locate urgent objective pain points, providing a concrete physical basis for product definition.

### Middle Layer (Theoretical Layer): Emotion–Value Segmentation Based on NLP

The theoretical layer focuses on users' psychological perception and emotional value structures. Even under identical physical environments, different cultural contexts lead to distinct interpretations of comfort. Natural language processing (NLP) techniques are applied to user-generated content on social media to analyze emotional orientations associated with comfort.

### Identification of Emotional Dimensions (Based on Positive / Negative Emotion Segmentation)

The analysis focuses on two primary emotional axes: positive and negative emotions. Positive emotional comfort emphasizes relaxation, security, freedom, and satisfaction, framing comfort as an actively pursued positive experience. In contrast, negative-oriented comfort frames comfort as relief from anxiety, fatigue, or stress, highlighting its function as psychological recovery, particularly in high-pressure social contexts. This emotional polarity reveals cultural differences in how comfort is psychologically constructed. (see Figure 2).



**Figure 2:** Emotion–value pathways based on NLP analysis.

### **Value–Semantic Association**

Further analysis can be conducted on high-frequency co-occurring words associated with “comfort.” If comfort-related terms frequently co-occur with words such as “family,” “gathering,” or “face,” it suggests that the culture is more strongly grounded in “social relationship harmony.” In such contexts, individual physical comfort often yields to social or ritual comfort. Conversely, if comfort-related terms frequently co-occur with words such as “being alone,” “privacy,” or “solitude,” it indicates that comfort is constructed around the establishment of personal boundaries. In this case, comfort is defined as successful defense against external interference.

### **Top Layer (Strategic Layer): Inferring Macro Life Paradigms and Strategies Based on the AEIOU Framework**

Strategic-layer research often relies on abstract policy texts; however, this study argues that macro-level national intentions and social structures ultimately materialize as concrete patterns of everyday life. To scientifically extract macro logic from micro phenomena, this study employs the AEIOU observation framework to structurally organize observed behaviors and environmental data (Hanington & Martin, 2019).

The AEIOU framework has been applied in empirical research. For example, Xiao and Hashim (2024) used AEIOU to analyze design requirements for kitchen tools for visually impaired users, extracting behavioral characteristics across five dimensions: activities, interactions, environments, objects, and users. In this study, the framework is elevated methodologically and treated as a “macro probe.” Combined with digital ethnography techniques, it analyzes common patterns across massive everyday-life samples to infer, from the bottom up, a country’s “survival strategy” and “cultural meta-logic.”

**Activities:** Activities reflect macro social rhythms and labor structures. Rather than focusing solely on observable actions, analysis emphasizes the systemic pressures underlying daily activities. When high-intensity, fast-paced activities dominate, this indicates an efficiency-oriented or high-pressure societal context, requiring comfort design to support dynamic regulation rather than static rest.

**Environments:** Environments serve as physical containers of resource constraints and urbanization strategies. Observations of spatial scale and infrastructural density reveal whether spatial efficiency constitutes a core survival logic. In highly compressed environments, comfort design must prioritize functional integration over formal spaciousness.

**Interactins:** Interactions reveal organizational patterns of social relationships and individual–collective boundaries. By examining interaction distance and dependency, it is possible to infer whether a society emphasizes collective interdependence or individual autonomy. Where independent interactions prevail, comfort design should focus on psychological coherence rather than social facilitation.

**Objects:** Objects indicate technological path dependencies in problem-solving. Analysis of the technological complexity of everyday objects reveals whether comfort is achieved through technological intervention

or environmental adaptation. In technology-dependent contexts, comfort enhancement should rely on digital augmentation and automation.

**Users:** Users reflect demographic structures and population trends. Observations of age and physical condition map broader issues such as aging or labor scarcity, requiring comfort-oriented design strategies to prioritize inclusivity, safety, and assistive functionality.

AEIOU Dimension	Micro-Behavioral Observations	Macro-Strategic Logic
A Activities	Action patterns, social norms, and daily behavioral rhythm.	Labor structure, social pace, and efficiency strategies.
E Environments	Physical scale, infrastructure traces, and resource conditions.	Resource endowment and urbanization/intensity strategies.
I Interactions	Social distance, frequency of contact, and dependency patterns.	Social organization paradigm (Collectivist vs. Atomized).
O Objects	Technical content of items and material/aesthetic attributes.	Industrial development level and technological path dependency.
U Users	Demographic features, body state, and subject characteristics.	National demographic trends (Aging, Labor shortage, etc.).

**Figure 3:** AEIOU mapping table.

Through systematic synthesis using the AEIOU framework (see Figure 3), it becomes possible to assemble a complete “macro lifestyle paradigm” from fragmented life slices. This inductive approach grounded in real scenarios reveals how macro forces concretely shape fundamental definitions of “what constitutes a comfortable life,” thereby providing the strategic layer with more explanatory empirical evidence than isolated statistical data.

## TRIANGULATION-BASED ANALYSIS AND VALIDATION

The core challenge of cross-regional user research lies in the invisibility and polysemy of “culture” itself. Different research perspectives often capture only partial facets of this phenomenon. Relying solely on macro statistical and policy data tends to reveal survival conditions and structural constraints, but makes it difficult to access emotional expectations and value orientations at the individual level. Focusing on emotional and semantic analysis helps to understand subjective experience, yet may weaken consideration of real material conditions and institutional environments. Meanwhile, behavior observation based on the AEIOU framework can present concrete forms of practice, but without higher-level analytical frameworks, it remains difficult to explain the underlying structural drivers and strategic logic.

To enhance the validity and reliability of research findings, this study introduces triangulation from the social sciences (Carter et al., 2014) as a systematic analytical strategy rather than a simple data-checking tool. Specifically, by cross-validating bottom-layer (material reality), middle-layer (psychological and emotional mechanisms), and top-layer (macro strategy and cultural structure) data, this study seeks to mitigate biases introduced

by single methods and to construct logical connections and explanatory closure across layers. The resulting design insights are thus grounded in objective survival conditions, responsive to subjective emotional needs, and interpretable within broader cultural and strategic contexts.

### **Validation Premises and Analytical Norms**

**Independent Coding:** To avoid “a priori” cognitive bias, data analysis across the three layers must remain logically independent. The research team should separately encode explicit data, social media corpora, and policy documents to derive preliminary conclusions for each layer, rather than allowing mutual influence at early stages of analysis.

**Semantic Alignment and Normalization:** Because the original forms of the three-layer data differ—such as statistical figures, emotional vocabulary, and behavioral patterns—it is necessary to translate them into a unified design semantic language.

## **CONVERGENT VALIDATION AND DIVERGENCE ANALYSIS**

### **Strong Convergence: Establishing Core Value Propositions**

When data from all three layers consistently point toward the same direction, this is defined as strong convergence. Such convergence indicates a highly reliable and non-negotiable user need.

For example, if environmental analysis reveals high noise levels, emotional analysis shows strong weighting of “quietness,” and behavioral observation identifies widespread use of noise-canceling solutions, auditory comfort can be identified as a core demand. In such cases, related features should be positioned as fundamental design priorities rather than optional enhancements.

### **Structural Divergence: Identifying Innovation Opportunities**

Inconsistencies or contradictions between layers do not necessarily indicate analytical failure. In cross-cultural contexts, such structural divergence often reflects latent cultural tension between idealized values and real-world conditions. Existing research emphasizes that the value of triangulation lies not in enforcing consistency, but in revealing deeper mechanisms through difference and conflict (Carter et al., 2014).

Therefore, relying solely on subjective expression may capture only surface-level desires, while examining inter-layer divergence enables more context-sensitive design interpretation.

### **Conflict Between Theoretical Layer and Practical Layer Data**

A representative form of tension emerges between idealized discourse and material reality. For instance, NLP-based analysis may show strong endorsement of “minimalism” and “decluttering,” while practical-layer data reveal small living spaces, high object density, and limited storage capacity.

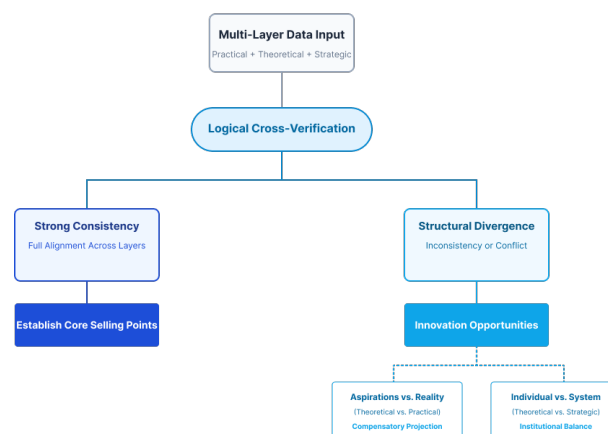
This contradiction reflects a compensatory value projection rather than genuine demand for functional reduction.

From a design perspective, such tension suggests that minimalist discourse should not be directly translated into reduced functionality. Instead, design should support perceived order and control without compromising actual usability.

### Conflict Between Theoretical Layer and Strategic Layer Data

Another significant tension arises between individual value preferences and macro-level governance strategies. Users may emphasize privacy and psychological security, while strategic-layer analysis reveals increasing data integration and algorithmic governance that limits individual control over information boundaries.

This tension reflects misalignment between value legitimacy and systemic rationality. In design practice, the goal is not to eliminate data collection, but to provide users with perceptible control and feedback, thereby reconstructing psychological security within existing institutional constraints.



**Figure 4:** Triangulation analysis process.

Through the above process (see Figure 4), this study is able not only to identify “what users want,” but also to delineate “what institutional boundaries allow” and “what real conditions support.” The resulting three-layer integrated analytical perspective effectively filters out pseudo-demands that appear attractive at the discursive or imaginative level but lack practical foundations, thereby enabling product definitions to achieve a more robust balance among commercial feasibility, technological feasibility.

## CONCLUSION

This paper addresses the academic challenge in cross-regional user research whereby macro-context analysis and micro-level experiential insights are fragmented from one another, and constructs a big-data-driven three-layer

analytical framework consisting of the practical layer, theoretical layer, and strategic layer. The core value of this methodology lies in its use of heterogeneous multi-source data to establish a translation mechanism from objective survival environments to subjective value logics.

In terms of the specific analytical pathway, the study integrates macro statistical scanning, NLP-based social semantic mining, and bottom-up inference of behavioral paradigms, enabling a systematic deconstruction of cross-cultural contexts.

On this basis, the significance of this research is mainly reflected in its support for the interpretation and application of cross-regional user research outcomes. By introducing a triangulation mechanism to differentiate and organize influencing factors across multiple layers, the research results can provide reference for determining the universality and contextual specificity of user needs in design practice. This helps designers make more reasonable experience trade-offs and design translations in cross-cultural scenarios.

## REFERENCES

- Bazzaz Abkenar, S., Haghi Kashani, M., Mahdipour, E., & Jameii, S. M. (2021). Big data analytics meets social media: A systematic review of techniques, open issues, and future directions. *Telematics and informatics*, 57, 101517.
- Bronfenbrenner, U. (1979). *The ecology of human development: Experiments by nature and design*. Harvard University Press.
- Bourdieu, P. (1977). *Outline of a theory of practice\**. Cambridge University Press.
- Carter, N., Bryant-Lukosius, D., DiCenso, A., Blythe, J., & Neville, A. J. (2014). The use of triangulation in qualitative research. *Oncology Nursing Forum*, 41(5), 545–547.
- Lee, J.-S., & Park, S. (2023). A cross-cultural anatomy of destination image: An application of mixed-methods of UGC and survey. *Tourism Management*, 98, 104746.
- Guo, Y., Guo, M., Su, J., Yang, Z., Zhu, M., Li, H., Qiu, M., Liu, S. (2024). Bias in Large Language Models: Origin, Evaluation, and Mitigation. arXiv preprint arXiv:2411.10915.
- Hanington, B., & Martin, B. (2019). *Universal methods of design: 125 ways to research complex problems, develop innovative ideas, and design effective solutions*. Rockport Publishers.
- Hong, D., Fu, Z., Zhang, X., & Pan, Y. (2025). Research on the Development and Application of the GDELT Event Database. *Data*, 10(10), 158.
- Karhapää, A., Pöysä-Tarhonen, J., Rikala, P., & Hämäläinen, R. (2025). Being an ethnographic researcher in a modern workplace: advantages and challenges of digital ethnography. *Journal of Organizational Ethnography*, 14(2), 233–250.
- Kleman, A., & Foster, S. (2025). The impact of apartment design on changes in housing satisfaction during the COVID-19 pandemic lockdown. *Cities & Health\**.
- Lee, J. (2025). From Big Data to Cultural Intelligence: An AI-Powered Framework and Machine Learning Validation for Global Marketing. *Journal of Theoretical and Applied Electronic Commerce Research*, 20(4), 288.
- None, D. J. R. (2025). Sentiment Analysis: Challenges and Insights. *Journal of Marketing & Social Research*, 2(8), 180–186.
- OECD. (2024). *OECD Economic Outlook, Volume 2024 Issue 1*. OECD Publishing.
- Reckwitz, A. (2002). Toward a theory of social practices: A development in culturalist theorizing. *European Journal of Social Theory*, 5\*(2), 243–263.

- Schein, E. H. (2010). *\*Organizational culture and leadership\** (4th ed.). Jossey-Bass.
- Wallerstein, I. (2004). *\*World-systems analysis: An introduction\**. Duke University Press.
- Wang, T. (2016). Big data needs thick data. In *The Ethos of Digital Ethnography* (pp. 25–38). Palgrave Macmillan.
- Wu, C., Zhu, T., Li, Y., Zhang, Z., & Wu, T. (2025). Balancing Multi-Source Heterogeneous User Requirement Information in Complex Product Design. *Symmetry*, 17(8), 1192.
- Wu, M., Mao, Y., Wu, S., Feng, J., Zhang, Q., Xie, T., Chen, H., & Zhu, T. (2023).
- Wu, S., Mao, Y., Wu, S., Feng, J., Zhang, Q., Xie, T., Chen, H. & Zhu, T. (2023). The big data analysis in cultural psychology. *Advances in Psychological Science*, 31(3), 317–329.
- Xiao, R., & Hashim, A. M. (2024). AEIOU analysis on tools design requirement for visually impaired users. *E-BPJ Proceedings*, 9(SI17), Article 5415.