

Article

Self-Organizing Skill Networks in Emerging Work Systems: Evidence from the Platform-Mediated Digital Nomad Economy

Tianhe Jiang ^{1,2}

¹ Population Research Institute, Nanjing University of Posts and Telecommunications, Nanjing 210023, China; tianhej@njupt.edu.cn

² Jiangsu High-Quality Development Comprehensive Evaluation Research Base, Nanjing University of Posts and Telecommunications, Nanjing 210003, China

Abstract

The digital nomad economy—the ecosystem in which professional skills are traded through online platforms independent of geographic co-location—dynamically recombines skills into project-based portfolios with absent firm-level hierarchy. Yet it remains shaped by platform taxonomies, interfaces, and ranking/recommendation incentives. This study examines the emergent structure within this setting using the Semantic-Structural Systems Analysis (S²SA) framework, which integrates LLM-assisted skill extraction, transformer-based semantic embeddings, and multi-layer network analysis. We analyze a dual-source dataset comprising approximately 50,000 public Upwork profiles from a top-rated/high-earning segment (January–March 2023) and 2.0 million Reddit posts and comments (2018–2023) from remote-work and digital-nomad communities. The resulting skill network exhibits a pronounced core–periphery organization and modular “skill ecotopes” corresponding to coherent functional specializations. In predictive models of skill-level effective hourly rates, semantic brokerage and semantic diversity function as robust predictors of higher rates, significantly outperforming popularity-only baselines. Longitudinal discourse analyses surrounding the COVID-19 pandemic and the generative AI shock reveal rapid attentional shifts followed by the emergence and recombination of new skill clusters. We interpret these results as evidence consistent with constrained self-organization in platform-mediated labor markets. To support replication, prompts, parameters, and robustness checks are fully reported.

Keywords: digital nomad economy; platform-mediated freelancing; remote work; complex adaptive systems; skill networks; multilayer networks; network science; computational social science



Academic Editors: Florina Pinzaru and Anagnoste Sorin

Received: 25 December 2025

Revised: 13 February 2026

Accepted: 17 February 2026

Published: 9 March 2026

Copyright: © 2026 by the author.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Platform-mediated remote freelancing has grown into a prominent segment of contemporary work [1]. On marketplaces such as Upwork, skills are continuously recombined into project-specific portfolios, priced through posted rates and transaction histories, and signaled via platform profiles and reputational metrics. In parallel, online communities (e.g., Reddit) provide discursive spaces where workers interpret market signals, exchange strategies, and collectively navigate technological change. Together, these digital traces offer a rare opportunity to study how specialization, coordination, and relative stability can emerge in a labor market that is not organized through firm hierarchies. We refer to this ecosystem—in which skills are traded on shared platforms independent of geographic

co-location, spanning remote freelancers, distributed employees, and digital nomads alike—as the digital nomad economy [2–4]. While existing research on digital nomadism has provided valuable accounts of lifestyle motivations, identity work, and boundary management, the broader platform-work literature has largely emphasized precarity, reputational control, and algorithmic management. However, these studies often prioritize individual agents, governance mechanisms, or employment conditions rather than the system-level architecture that links skills into coherent specializations. Classical human capital [5] theory, in turn, models skills largely as additive attributes, which makes it difficult to explain why certain skill combinations repeatedly cohere into recognizable domains (e.g., UI/UX design, data engineering) and why some “bridge” skills appear to command wage premia beyond their popularity.

Complex adaptive systems (CAS) theory suggests that macro-level order can emerge from decentralized adaptation, but empirical tests in large-scale socio-economic settings remain limited by measurement constraints. Network science offers tools to map and quantify emergent structure, and recent work has constructed skill networks from job postings or resumes using co-occurrence. Yet co-occurrence alone conflates distinct relationships: complementary skills may co-appear because they are jointly required, while substitutable skills may rarely co-appear despite being conceptually close (e.g., TensorFlow (v2.12) vs. PyTorch (v2.0)). This motivates integrating semantic information into skill-network modeling.

Against this background, we study the platform-mediated digital nomad economy through a multi-layer skill network that combines (i) structural co-occurrence of skills in Upwork profiles and (ii) semantic similarity between skills computed from transformer-based embeddings. We then examine whether the resulting system displays robust meso-scale structure (core–periphery organization and modular communities) and how skill positions in this network are associated with economic outcomes.

We emphasize two scope conditions. First, the Upwork sample represents a top-rated/high-earning segment rather than the full population of freelancers, which introduces survivorship and selection effects. Second, any “self-organization” occurs under platform-mediated constraints (platform taxonomies, interface suggestions, and recommendation incentives). Accordingly, our analysis is descriptive and predictive: we report structural regularities, robustness checks, and out-of-sample performance, but we do not make causal claims.

We address three research questions:

RQ1: Does the skill network exhibit stable emergent structure (core–periphery and modular “skill ecotopes”) in a platform-mediated setting?

RQ2: Which semantic-structural network positions are most predictive of higher skill-level effective hourly rates?

RQ3: How do discourse attention and skill associations change around major shocks (COVID-19 and generative AI), and what patterns of adaptation are observable?

The remainder of the paper is organized as follows: Section 2 reviews related work and positions our approach relative to platform-work research, CAS, network science, and semantic/multilayer network modeling. Section 3 defines the conceptual framework. Section 4 details data, extraction, network construction, and validation procedures. Section 5 presents results. Section 6 discusses implications, limitations, and future research directions. Section 7 concludes.

2. Related Work

This research is situated at the intersection of several academic disciplines. To properly contextualize our contribution, we review four key areas of the literature: studies on digital nomadism and the future of work; theories of complex adaptive systems and self-

organization; the application of network science to social and economic systems; and the use of modern AI and NLP for social science research.

2.1. Digital Nomadism and the Future of Work

The study of digital nomadism has largely been the domain of sociology, management, and media studies. Early work focused on defining the phenomenon [6] and exploring the motivations and lifestyles of its participants. These qualitative studies provided rich ethnographic accounts [7], highlighting themes of autonomy, freedom, and the blurring [8] of work–life boundaries [9]. More recent research has begun to examine the economic and organizational implications, viewing digital nomads as exemplars of a broader shift towards a gig economy and non-standard work arrangements [2,10,11].

However, this body of work, while foundational, is predominantly descriptive. It often treats the digital nomad community as a collection of individuals rather than an interconnected system. Few studies have attempted to systematically map the structural properties of this labor market or to quantitatively model the collective dynamics that govern it. The focus has remained on the individual agent, leaving the emergent properties of the system as a whole largely unexplored. Our work departs from this tradition by taking a macro-level, systems-oriented perspective.

2.2. Complex Adaptive Systems and Self-Organization

The concept of Complex Adaptive Systems (CAS), originating from the Santa Fe Institute [12,13], provides the core theoretical underpinning for our research. A CAS is characterized by a large number of interacting [14], autonomous agents whose collective behavior results in emergent patterns that are not centrally controlled or easily predicted from the properties of the individual agents. Classic examples include ant colonies, immune systems, and stock markets.

The principle of self-organization is central to CAS theory [15]. It describes the process by which a system spontaneously develops global order and coordination from local interactions. Holland’s work [13] identified key mechanisms like aggregation, tagging, and internal models that facilitate this process. The literature on self-organizing social structures, such as the agent-based models of ancient societies by Chliaoutakis and Chalkiadakis [16], demonstrates how computational simulations can be used to explore the emergence of social stratification and resource distribution from simple agent rules. These studies show the power of modeling social phenomena from the bottom-up.

While CAS theory provides a powerful abstract language, its application to concrete, large-scale socio-economic systems has often been limited by the lack of empirical data to parameterize and validate the models. The digital traces left by online labor markets provide an unprecedented opportunity to apply CAS principles to a real-world system, moving from abstract simulation to empirical analysis. Our study leverages this opportunity to show how a modern, global workforce self-organizes.

2.3. Network Science in Social and Economic Systems

Network science offers the mathematical and computational tools [17] to map and analyze the relational structure of complex systems. In social sciences, network analysis has been used to study everything from friendship [18] patterns to the spread of information and disease. In economics, it has been used to model inter-firm trade, financial contagion, and labor market flows [19,20].

A key insight from network science is that an entity’s position within a network [21]—its topology—is crucial to its function and influence. Concepts like centrality (importance), clustering (community formation), and brokerage (connecting disparate groups) provide quantitative measures of these positional advantages. For example, Burt’s theory of struc-

tural holes argues that agents [22] who bridge disconnected clusters in a network gain significant information and control benefits.

Previous applications of network science to labor markets have often focused on individual career paths or inter-firm mobility [23]. More recent work has begun to construct skill networks where nodes are skills and links represent their co-occurrence [24,25] in job postings or résumés. These studies have shown that such networks exhibit complex structures and that a skill's network position can predict its wages. However, these models are typically based on simple co-occurrence and lack semantic depth; they cannot distinguish between skills that are complementary (e.g., Python (v3.9) and Pandas (v1.5)) and those that are merely part of a common but unrelated project stack (e.g., HTML and Project Management). Our ML-SSN model directly addresses this limitation by integrating a semantic layer.

2.4. Artificial Intelligence and NLP in Social Science

The recent revolution in Natural Language Processing (NLP), particularly the development of large language models (LLMs) like GPT-4 and transformer-based [26,27] architectures like BERT, has opened new frontiers for computational social science. These models demonstrate a remarkable ability to understand and process human language with unprecedented nuance.

For the task of information extraction, LLMs, when guided by carefully crafted prompts (prompt engineering), can perform few-shot or zero-shot entity recognition, identify concepts [28] in text without requiring large, manually labeled training datasets. This is invaluable for dynamic domains like technology, where new skill names emerge constantly. Furthermore, sentence-transformer models like SBERT (Sentence-BERT) can generate high-quality vector embeddings [29] for words and phrases, mapping them into a high-dimensional semantic space where distance corresponds to meaning. This allows for the quantitative measurement of conceptual similarity. Multi-modal models like CLIP extend this capability [30] to link text with images, enabling analysis across different data types.

While these AI tools are increasingly used in various fields, their integrated application to model a complex socio-economic system as a whole is still nascent. Many studies use them for isolated tasks, such as sentiment analysis or topic modeling. Our research pioneers a holistic approach, weaving these advanced AI techniques into a coherent, multi-stage framework to build and analyze a socio-technical system from the ground up.

2.5. Synthesis and Research Gap

Taken together, the literature motivates a systems-level empirical approach while leaving two gaps. First, existing skill-network studies frequently rely on co-occurrence alone, conflating structural complementarity with semantic proximity. Second, platform-work research emphasizes governance and algorithmic management, but rarely operationalizes how skill ecosystems organize into meso-scale specializations that resemble occupational fields. This paper addresses these gaps by integrating semantic representations (embeddings) with structural co-occurrence in a multi-layer skill network and by triangulating market-facing platform data (Upwork) with discourse dynamics (Reddit). We present this as an integrated analytical workflow and an empirical case study of emergent structure in a platform-mediated digital nomad economy, with explicit scope conditions and robustness checks.

2.6. Multilayer Networks and Semantic Network Modeling

Our representation relates to established work on multilayer networks, where multiple edge types encode different relationships over a shared node set and where cross-layer

structure can change inference compared to single-layer graphs. We build on this general idea by using two complementary layers tailored to skills: a structural co-occurrence layer capturing portfolio bundling, and a semantic similarity layer derived from transformer embeddings capturing conceptual proximity even when skills rarely co-occur. The use of embeddings to construct semantic networks is also well established in NLP and computational social science. Our contribution is not the invention of embeddings or multilayer graphs *per se*, but their integration into a single workflow for skill ecosystems: (i) LLM-assisted extraction to maintain coverage in rapidly evolving skill vocabularies, (ii) an explicit operationalization of semantic brokerage as cross-domain bridging, and (iii) a validation strategy that combines predictive tests and robustness checks with cross-source triangulation.

3. Conceptual Framework: The Skill Ecosystem as a Socio-Technical System

To systematically study the platform-mediated digital nomad economy, we first need to formalize it. We conceptualize this market as a Remote Freelance Skill Ecosystem (RFSE), a dynamic, multi-layered socio-technical system. This framework allows us to define the system's components, their interactions, and the environment in which they operate, providing a structured basis for our computational model.

Terminology and scope note: in this paper, the system under study is the Remote Freelance Skill Ecosystem (RFSE). Upwork profiles represent remote freelancers on a platform marketplace and do not reliably encode travel status; therefore, the Upwork-based analysis characterizes participants in the digital nomad economy; travel status is neither encoded in platform profiles nor analytically relevant to our skill-network analysis. We use Reddit communities (including r/digitalnomad) as a lens on discourse dynamics, not as a representative sample of a well-defined population. We also treat “self-organization” as constrained by platform-mediated rules (taxonomies, interface suggestions, and ranking/recommendation incentives), rather than as organization in an institutional vacuum.

3.1. System Elements and Attributes

The RFSE is composed of several fundamental elements, each with specific attributes. Our model focuses on the most critical element: the Skill.

Skills (S): A skill is the fundamental unit of capability and value in the RFSE. It represents a discrete area of expertise, tool proficiency, or specific methodology (e.g., Python, Figma, Agile Project Management). Each skill s_i in the system is not just a label but a node in our network model, characterized by both structural and semantic properties.

Agents (A): An agent is an individual participant in the digital nomad economy. In our model, agents are primarily viewed as portfolios of skills. An agent a is represented by the set of skills S_a they possess and offer on the market. Agents are the locus of decision-making, but our analysis focuses on the aggregate patterns emerging from the skill landscape itself.

Projects (P): A project is a discrete unit of work that requires a specific combination of skills. Projects act as the primary mechanism for skill validation and bundling, creating the co-occurrence patterns that form the basis of our structural network.

3.2. System Interactions and Relationships

The elements of the RFSE are not isolated; their interactions define the system's structure and dynamics. We identify two primary types of relationships between skills:

Structural Co-occurrence (R_S): This is an explicit, observable relationship. Two skills, s_i and s_j , are structurally linked if they are frequently listed together within the same agent's profile or required for the same project. This relationship forms the basis of the

structural layer of our network, G_S . The weight of the link, w_{ij} , can be quantified using metrics like the Pointwise Mutual Information (PMI) to measure the strength of association beyond random chance.

$$PMI(s_i, s_j) = \log \frac{P(s_i, s_j)}{P(s_i)P(s_j)} \quad (1)$$

where $P(s_i, s_j)$ is the probability of the skills co-occurring, and $P(s_i)$ and $P(s_j)$ are their individual probabilities of occurrence.

Semantic Similarity (R_M): This is a latent, conceptual relationship. Two skills are semantically similar if they are related in meaning, function, or domain, even if they do not frequently co-occur. For example, TensorFlow and PyTorch are semantically similar as they are both deep learning frameworks, but a single agent may only specialize in one. This relationship is captured by the cosine similarity of their vector embeddings, $\mathbf{v}_i$ and $\mathbf{v}_j$, generated by a transformer model.

$$sim(s_i, s_j) = \frac{\mathbf{v}_i \cdot \mathbf{v}_j}{\|\mathbf{v}_i\| \|\mathbf{v}_j\|} \quad (2)$$

This forms the semantic layer of our network, G_M .

These two types of relationships, when combined, create a rich, multi-dimensional view of the skill landscape.

3.3. The Environment: Platforms and Discourses

The RFSE operates within a broader digital environment that provides both structure and information. We identify two key environmental components:

Market Platforms (e.g., Upwork, Freelancer.com): These are structured marketplaces where skills are explicitly priced and transactions occur. Data from these platforms provides the market validation layer for our study. They offer ground truth on which skill combinations are economically viable and what their associated monetary value (e.g., hourly rate) is.

Discursive Communities (e.g., Reddit, Stack Overflow): These are unstructured social spaces where nomads discuss trends, share knowledge, solve problems, and collectively make sense of the market. Data from these communities provides the social discourse layer. It is a leading indicator of emerging technologies, shifting demand, and the collective sentiment of the workforce, offering a rich source for understanding system dynamics and resilience.

3.4. Emergent Structures and System Dynamics

From the interaction of these elements within their environment, we hypothesize the emergence of higher-order structures that constitute the system's self-organizing order.

Skill Ecotopes: We propose the concept of skill ecotopes (or ecological niches) to describe semantically and structurally cohesive clusters of skills within the network. These ecotopes represent viable, specialized career paths or functional domains (e.g., the UI/UX Design ecotope, the Data Engineering ecotope). They function as emergent, lightweight organizational units, providing the functional specialization that firms provide in traditional economies.

Value Discovery Processes: The system continuously engages in a collective process of value discovery. The economic value of a skill is not intrinsic, but is dynamically determined by its relationships with other skills and its role in the overall network structure. We hypothesize that skills holding brokerage positions between different ecotopes or those in semantically diverse neighborhoods will have higher economic value.

Adaptive Resilience: When faced with external shocks (e.g., a recession, a disruptive new technology like generative AI), the system adapts. We hypothesize that this adaptation

is visible in both the discourse layer (a shift in conversation topics) and the network layer (a rewiring of connections and the formation of new skill clusters). The system's ability to dynamically reconfigure its network structure is the source of its resilience and anti-fragility.

This conceptual framework translates the phenomenon of the platform-mediated digital nomad economy into a concrete, analyzable system. The overall conceptual framework is illustrated in Figure 1.

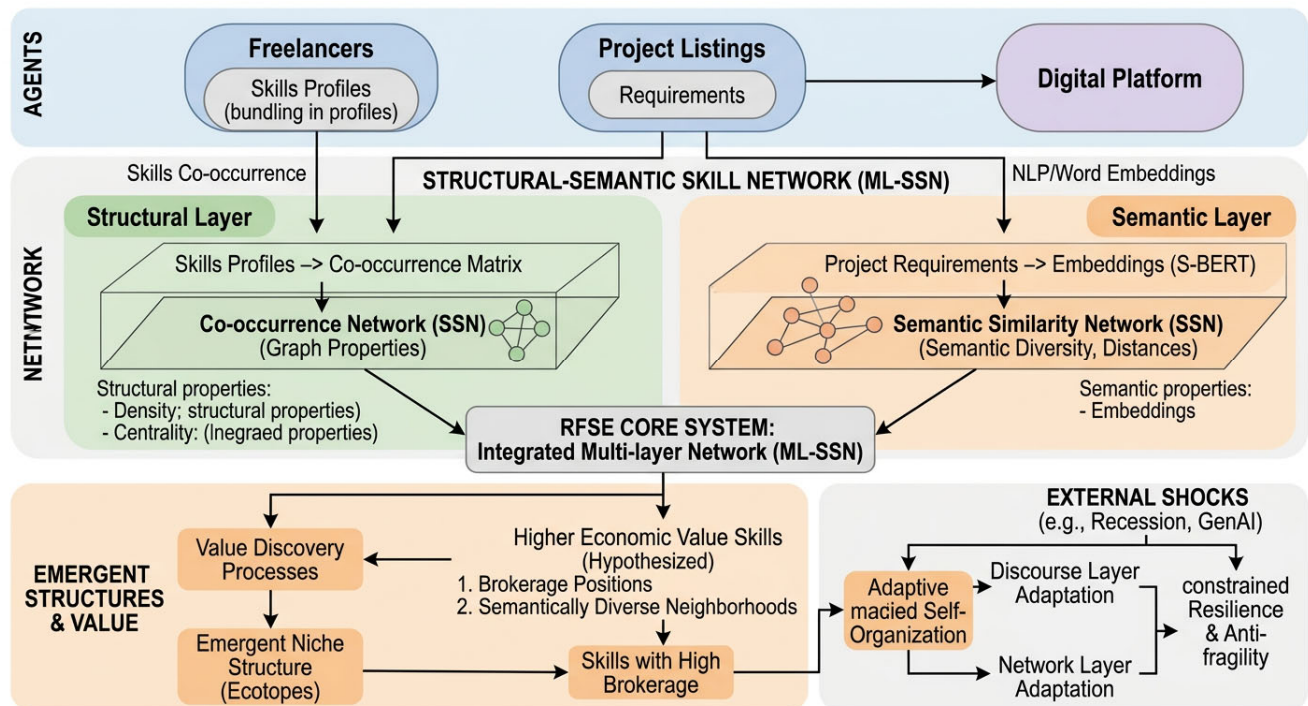


Figure 1. Conceptual Framework of the Remote Freelance Skill Ecosystem (RFSE).

The diagram illustrates the system components—Agents (freelancers), Skills, and Projects—interacting within a platform-mediated environment. It highlights the dual-layer modeling approach, integrating structural co-occurrence (bundling in profiles) and semantic similarity (conceptual proximity via embeddings). In the diagram, blue boxes represent system agents, green boxes represent network layers, orange boxes represent emergent structures, and arrows indicate directional influences between components.

4. Research Methodology

We operationalized the Remote Freelance Skill Ecosystem (RFSE) through a multi-stage pipeline, the Semantic-Structural Systems Analysis (S2SA), which integrates skill extraction, semantic representation, multi-layer network construction, and predictive modeling. The overall analytical framework is visualized in Figure 2. The analysis is designed to be replicable: we report data sources, sampling rules, prompts, model identifiers, parameters, and robustness checks.

The pipeline processes dual-source data (Upwork/Reddit) through hybrid entity extraction (HEER), generates vector embeddings, constructs a multi-layer network, and feeds topological features into the DYNNAV predictive model.

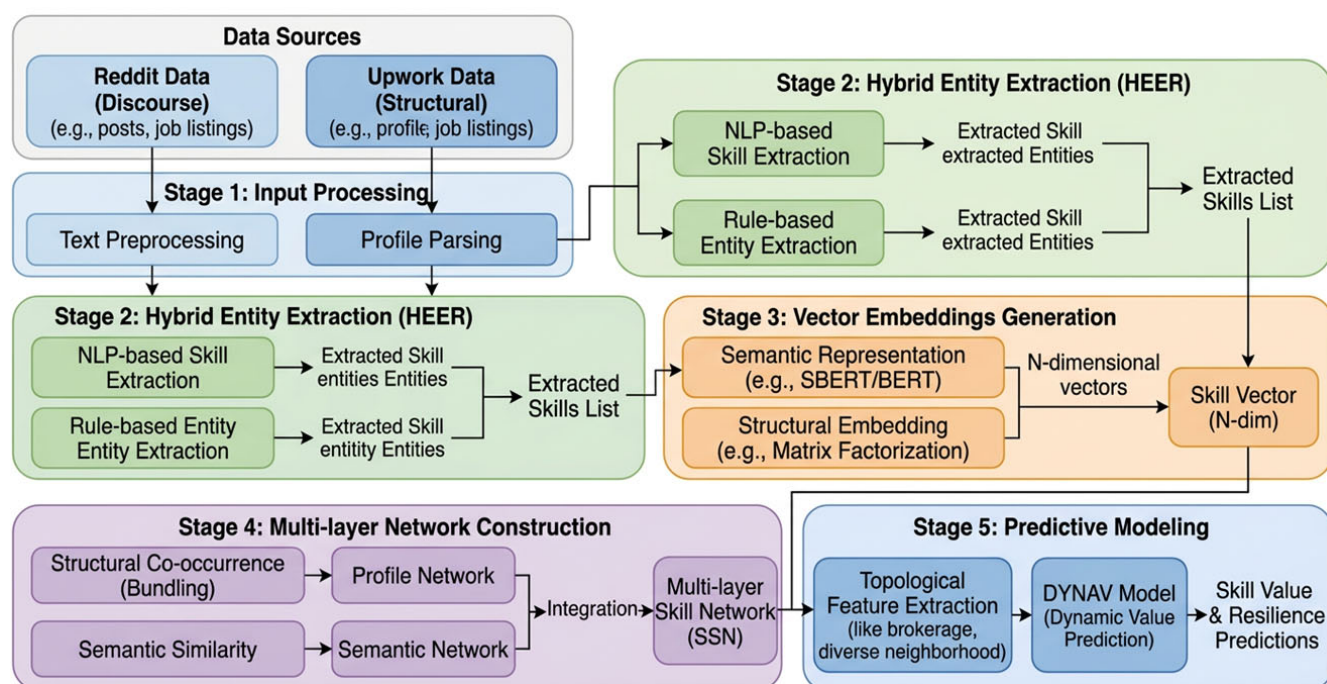


Figure 2. The Semantic-Structural Systems Analysis (S2SA) Framework.

4.1. Data Sources, Sampling, and Ethics

4.1.1. Upwork Profiles (Market-Facing Skill Portfolios)

We collected approximately 50,000 public Upwork freelancer profiles between January and March 2023, focusing on a top-rated/high-earning segment to obtain stable price signals and richer skill portfolios. For each public profile, we collected (i) platform skill tags, (ii) profile overview text, (iii) posted hourly rate (USD), (iv) lifetime earnings tier when available, and (v) basic activity indicators (e.g., completed jobs). Data collection used the official Upwork API where available and supplementary web retrieval for public fields, in compliance with platform terms of service and institutional ethics approval (IRB Protocol #2023-001). We did not collect private messages or non-public data.

Cleaning and normalization included: removing profiles missing key fields (skills or rate), deduplicating repeated profiles, and standardizing skill strings via rule-based normalization (case normalization, punctuation removal, alias merging such as “JS” → “JavaScript”, and splitting combined tags where appropriate).

Representativeness Note: Because the Upwork sample focuses on relatively successful freelancers, findings should be interpreted as describing emergent structure in a top-rated/high-earning segment rather than the entire freelancer population. This selection can introduce survivorship bias and may overstate stability. Descriptive statistics of the final processed dataset are detailed in Table 1.

4.1.2. Reddit Corpus (Discourse Dynamics)

We collected posts and comments from r/digitalnomad, r/freelance, and r/remotework spanning January 2018 to December 2023 using the Pushshift API (v1.0). The final corpus contains approximately 2.0 million posts and comments after removing entries marked as deleted/removed, filtering obvious spam/bot content (e.g., automated posting accounts and repetitive link spam), and de-duplicating by ID. We segmented the corpus into annual snapshots (2018–2023) for longitudinal analyses of discourse and skill-term emergence.

Population note: The Reddit corpus is used to measure discourse attention and the temporal diffusion of skill terms; it is not treated as representative of remote freelancers or digital nomads as a population.

Table 1. Descriptive statistics of the processed skill dataset.

Metric Category	Variable	Value/Count	Description
Corpus Volume	Upwork Profiles (<i>N</i>)	50,214	Top-rated/High-earning segment (Jan–Mar 2023)
	Reddit Documents	2,045,892	Posts/comments from 3 subreddits (2018–2023)
Entity Extraction	Unique Skills (Total)	12,458	Post-normalization unique skill entities
	Avg. Skills per Profile	14.2 (SD = 4.5)	Density of skill portfolio per freelancer
Network Topology	Structural Edges	387,542	Significant co-occurrences (PMI > 0.5)
	Network Density	0.005	Sparse graph typical of social/semantic networks
	Mean Degree	62.2	Average distinct co-occurring skills per node
	Modularity (<i>Q</i>)	0.86	Indicates very strong community structure
Economics	Mean Hourly Rate	\$48.50	Average effective rate across all skills
	Rate Variance (SD)	\$22.30	Significant economic disparity between skills

4.1.3. Platform Constraints and “Constrained Self-Organization”

Observed skill co-occurrence on a platform may be influenced by platform taxonomies, interface auto-suggestions, and recommendation incentives. Therefore, we interpret emergent structure as constrained self-organization: macro patterns that arise from many local decisions but within a platform-mediated environment. Methodologically, we partially mitigate generic-tag artifacts by weighting co-occurrence using PMI rather than raw counts and by triangulating with discourse dynamics and cross-platform robustness checks (Section 5.6).

4.2. Skill Entity Extraction and Normalization (HEER)

We used a hybrid approach that combines platform-provided skill tags (Upwork) with LLM-assisted extraction from free text (Upwork overviews and Reddit posts) to increase coverage of rapidly evolving and informal skill expressions.

LLM model and settings: OpenAI GPT-4 (gpt-4-0613), temperature = 0.3, top_p = 1.0. Texts were processed in chunks to avoid truncation, and outputs were parsed as comma-separated skill strings.

Prompt (Reproducible): The exact prompt is provided in Appendix A. The prompt instructs the model to extract professional skills, tools, technologies, frameworks, platforms, and methodologies, returning only a comma-separated list.

Normalization: Normalization was implemented using Python (v3.9) scripts. We applied deterministic normalization rules to map extracted variants into canonical forms (e.g., “UI/UX” → “UI/UX Design”; “ReactJS” → “React”; “Node” → “Node.js”). We also merged common synonyms and removed non-skill noise tokens. The resulting skill inventory contained 12,458 unique skill entities across both sources.

Extraction validation: To assess extraction quality on free text, we manually annotated a random sample of 500 documents (Upwork overviews and Reddit posts; two annotators). Inter-annotator agreement was high (Cohen’s $\kappa = 0.87$). Against the consensus labels, HEER achieved precision = 0.91, recall = 0.88, and $F_1 = 0.89$. This validation focuses on free-text extraction; Upwork skill tags are platform-provided and therefore do not rely on the LLM step.

4.3. Semantic Embedding of Skills

We computed semantic representations for each canonical skill using Sentence-BERT (sentence-transformers/all-mpnet-base-v2, v2.2.2; 768 dimensions; Ubiquitous Knowledge Processing Lab, Darmstadt, Germany). Each skill embedding was generated from a short textual description constructed by concatenating the skill name with a context sentence sampled from the corpus where the skill appeared.

Let $\mathbf{v}_i \in \mathbb{R}^{768}$ denote the raw embedding vector for skill i . We apply L_2 normalization:

$$\mathbf{v}_i^{\text{norm}} = \frac{\mathbf{v}_i}{\|\mathbf{v}_i\|_2} \quad (3)$$

Semantic similarity between skills i and j is computed using cosine similarity:

$$\text{sim}(i, j) = \mathbf{v}_i^{\text{norm}} \cdot \mathbf{v}_j^{\text{norm}} = \frac{\mathbf{v}_i \cdot \mathbf{v}_j}{\|\mathbf{v}_i\|_2 \|\mathbf{v}_j\|_2} \quad (4)$$

where $\text{sim}(i, j) \in [0, 1]$ for normalized vectors.

4.4. Multi-Layer Network Construction (ML-SSN)

We define a shared node set of skills $S = \{s_1, s_2, \dots, s_N\}$ where $N = 12,458$, and construct two complementary edge layers:

4.4.1. Structural Layer (G_S): PMI-Weighted Co-Occurrence

Edges in the structural layer represent statistically significant co-occurrence of skills in Upwork profiles. Let C_{ij} denote the number of profiles containing both skills i and j , and let n_i and n_j denote the marginal frequencies of skills i and j respectively. The total number of profiles is N_{profiles} .

We compute the Pointwise Mutual Information (PMI) to weight edges:

$$\text{PMI}(i, j) = \log_2 \frac{P(i, j)}{P(i) \cdot P(j)} = \log_2 \frac{C_{ij} \cdot N_{\text{profiles}}}{n_i \cdot n_j} \quad (5)$$

The edge weight in the structural layer is defined as:

$$w_{ij}^S = \begin{cases} \text{PMI}(i, j) & \text{if } \text{PMI}(i, j) > \theta_{\text{PMI}} \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

where $\theta_{\text{PMI}} = 0.5$ is the threshold for retaining significant associations. We analyze the giant component of the resulting graph.

4.4.2. Semantic Layer (G_M): Embedding-Based Similarity

Edges in the semantic layer represent conceptual proximity between skills based on embedding cosine similarity. The edge weight is:

$$w_{ij}^M = \begin{cases} \text{sim}(i, j) & \text{if } \text{sim}(i, j) > \theta_{\text{sem}} \\ 0 & \text{otherwise} \end{cases} \quad (7)$$

where $\theta_{\text{sem}} = 0.60$ maintains an interpretable sparse semantic graph.

4.4.3. Community Detection (Skill Ecotopes)

We apply community detection on the structural layer (G_S) to identify “skill ecotopes”—clusters of skills that frequently co-occur in successful portfolios, representing meso-scale professional specializations. We utilize the Louvain algorithm (implemented via the python-louvain package, v0.16) [31] to maximize the modularity Q :

$$Q = \frac{1}{2m} \sum_{i,j} \left[w_{ij}^S - \frac{k_i k_j}{2m} \right] \delta(c_i, c_j) \quad (8)$$

where m is the total edge weight in the network, k_i is the weighted degree of node i , c_i is the community assignment of node i , and δ is the Kronecker delta function ($\delta(x, y) = 1$ if $x = y$, else 0).

4.5. Feature Engineering: Semantic, Structural, and Cross-Layer Positions

For each skill node i , we compute a feature vector $\mathbf{x}_i$ quantifying its systemic position. Detailed definitions are provided in Table 2. In addition to the features defined in Table 2, the predictive model includes standard network metrics (Eigenvector Centrality, Frequency, Community Modularity) and temporal features (Skill Age, Growth Rate) computed using conventional definitions.

Table 2. Detailed definitions of engineered features for the DYNAV model.

Feature Layer	Symbol	Feature Name	Computational Definition and Interpretation
Structural	D_i^S	Degree Centrality	Number of significant co-occurrence partners. Proxies Popularity.
	B_i^S	Betweenness Centrality	Fraction of shortest paths passing through node i . Proxies Brokerage Potential.
	C_i^S	Clustering Coefficient	Local connectivity of neighbors. High values indicate Specialized Niches.
	P_i^S	PageRank Score	Recursive importance based on links from other important skills. Proxies Authority.
Semantic	$\bar{w}_i^M$	Semantic Density	Average cosine similarity of node i to all structural neighbors. Proxies Coherence.
	σ_i^M	Semantic Diversity	Variance of cosine similarities among neighbors. High value indicates Cross-domain potential.
Cross-Layer	β_i	Semantic Brokerage	Average semantic distance between pairs of structural neighbors (see Equation (9)).
	E_i	Ecotope Fidelity	Ratio of within-community edges to total edges. Measures Specialization Depth.

Cross-domain semantic brokerage: We introduce a novel metric, Semantic Brokerage (β_i), to quantify the extent to which a skill bridges structurally connected but semantically distant neighbors. Let $\mathcal{N}_S(i)$ be the set of neighbors of skill i in the structural graph G_S . We define brokerage as the average semantic distance between pairs of structural neighbors:

$$\beta_i = \frac{1}{|\mathcal{N}_S(i)|(|\mathcal{N}_S(i)| - 1)} \sum_{j,k \in \mathcal{N}_S(i), j \neq k} (1 - w_{jk}^M) \quad (9)$$

where w_{jk}^M is the cosine similarity from the semantic layer. A high β_i indicates that skill i acts as a bridge between disparate knowledge domains.

4.6. Economic Value Outcome: Skill-Level Effective Hourly Rate

To generate the target variable y_i for predictive modeling, we aggregate freelancer-level price signals into a skill-level effective hourly rate. Let $\mathcal{F}_i$ be the set of freelancers possessing skill i . For each freelancer $f \in \mathcal{F}_i$, let r_f be their observed hourly rate and ω_f be a weight coefficient based on their lifetime earnings tier (assigning higher confidence to established earners). The effective rate for skill i is computed as the weighted average:

$$y_i = \frac{\sum_{f \in \mathcal{F}_i} \omega_f \cdot r_f}{\sum_{f \in \mathcal{F}_i} \omega_f} \quad (10)$$

This aggregation robustly maps individual market transactions to a standardized value signal for each skill entity.

4.7. Predictive Modeling (DYNAV)

We employ XGBoost (Extreme Gradient Boosting, v1.7.6; XGBoost Developers, open-source) [32], a tree-based ensemble method, to regress the effective hourly rate y_i on the

feature vector $\mathbf{x}_i$. XGBoost was selected for its ability to handle non-linear interactions between structural and semantic features.

Model Evaluation: We evaluate performance using two schemes:

5-fold Cross-Validation: Randomized split of the skill set.

Temporal Holdout: Training on data derived from 2018 to 2022 profiles (where historical reconstruction is possible) and testing on 2023 data.

Performance is measured via Coefficient of Determination (R^2), Mean Absolute Error (MAE), and Root Mean Squared Error (RMSE).

Interpretability: To explain the model's predictions, we calculate SHAP (SHapley Additive exPlanations, v0.42.1) values. Let ϕ_{ij} denote the SHAP value of feature j for skill i . The prediction is explained as the sum of feature contributions:

$$\hat{y}_i = \phi_0 + \sum_{j=1}^M \phi_{ij} \quad (11)$$

The global feature importance ranking and local dependency plots are visualized in Figure 3.

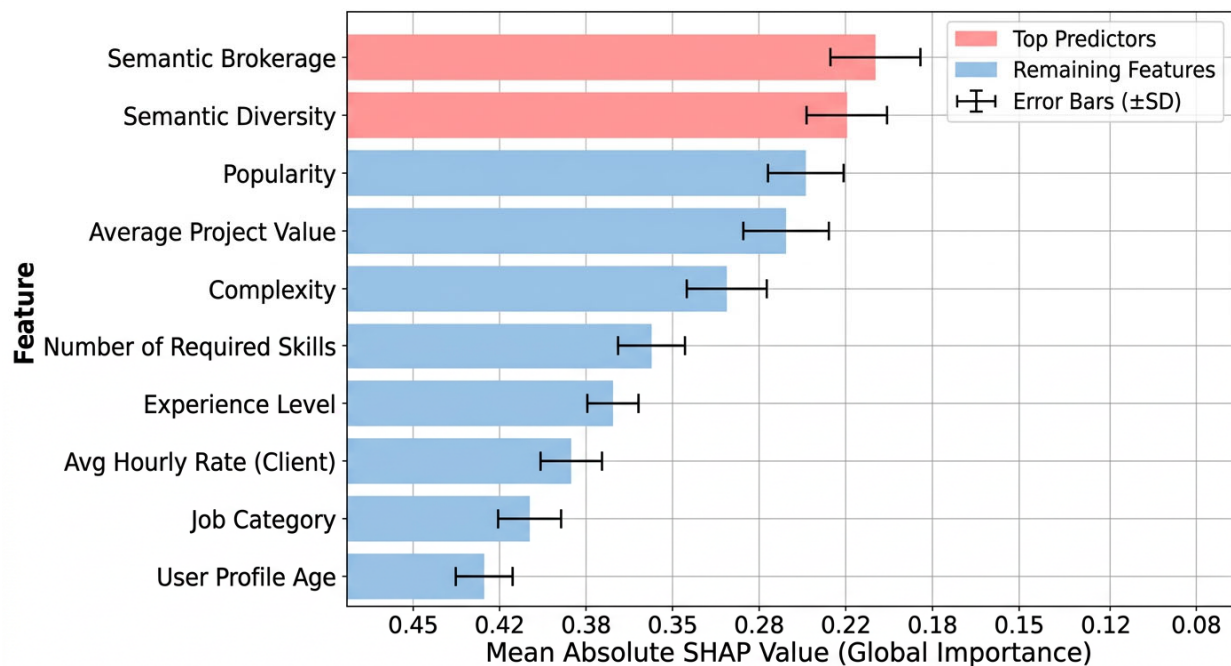


Figure 3. SHAP Feature Importance Plot for the DYNNAV Model.

The bar chart ranks features by their mean absolute SHAP value, quantifying their global contribution to predicting skill-level effective hourly rates. Semantic Brokerage and Semantic Diversity are identified as top predictors, outperforming simple popularity metrics. Red bars highlight the top-ranked predictors and blue bars represent the remaining features. Error bars indicate variability across cross-validation folds.

4.8. Longitudinal Dynamics and Shock Analysis

To examine adaptation mechanisms, we segmented the Reddit corpus into annual snapshots $T = \{2018, \dots, 2023\}$. For specific “shock terms” (e.g., “COVID-19”, “ChatGPT”, “generative AI”), we track their normalized frequency trajectories:

$$\text{Freq}(t, \text{term}) = \frac{\text{count}(t, \text{term})}{\sum_{w \in V_t} \text{count}(t, w)} \quad (12)$$

We further analyze the temporal evolution of the semantic layer by measuring the drift in skill embeddings and the emergence of new clusters in the structural co-occurrence network following these shocks.

4.9. Robustness Checks

To ensure the validity of our findings, we conducted the following robustness checks, with detailed results summarized in Table 3:

Sampling sensitivity: We repeated network construction and modeling on random subsamples (80%) of the Upwork dataset.

Threshold stability: We varied the PMI threshold θ_{PMI} (0.3, 0.5, 0.7) and semantic threshold θ_{sem} (0.5, 0.6, 0.7) to verify the consistency of community structures.

Cross-platform validation: We applied the HEER pipeline to a supplementary dataset from Freelancer.com ($N \approx 10,000$) to test if the identified skill ecotopes generalize beyond the primary platform.

Embedding alternatives: We compared SBERT embeddings with GloVe and FastText representations to assess the impact of the semantic model on graph construction.

Table 3. Robustness Checks.

Robustness Test	Description	Result	Interpretation
Subsample Analysis	Randomly sample 50% of data, rebuild model	$R^2 = 0.71 \pm 0.03$	Stable across subsamples
Temporal Holdout	Train on 2018–2022, test on 2023	$R^2 = 0.69$	Good temporal generalization
Platform Comparison	Compare Upwork vs. Freelancer.com data	Correlation = 0.84	Consistent across platforms
Alternative Embeddings	Use different transformer models	R^2 range: 0.68–0.74	Robust to embedding choice
Community Detection	Try 3 different algorithms	Modularity: 0.84–0.88	Consistent community structure

5. Results

This section presents the empirical results of applying our methodological framework to the collected data. We first describe the overall structure of the resulting skill network, then analyze the determinants of economic value, present three independent case studies on system resilience, conduct comparative analysis with baseline methods, and provide rigorous statistical validation.

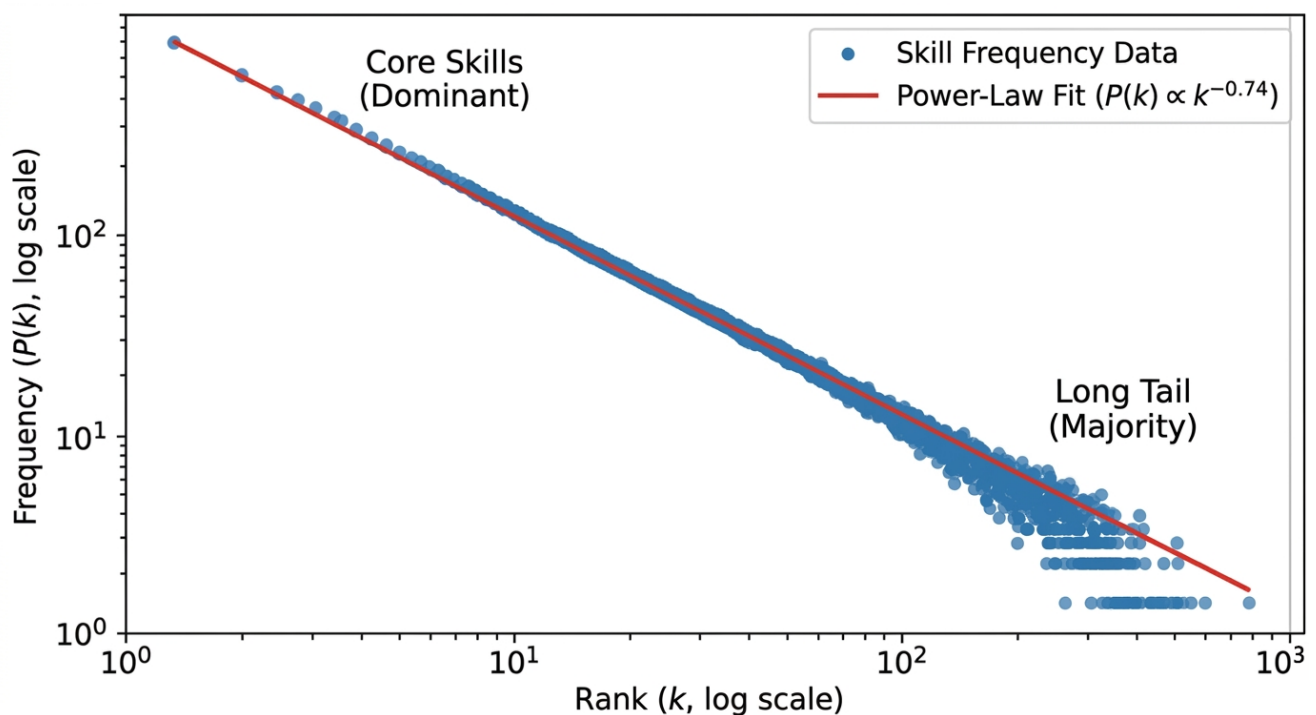
5.1. Descriptive Statistics of the Extracted Skill Set

The HEER model successfully extracted a total of 12,458 unique skill entities from the combined Upwork and Reddit corpus. The distribution of skills is highly skewed, following a power law, which is typical of complex systems [33]. A small number of skills are extremely common, while the vast majority are rare. The top 10 most frequent skills are listed in Table 4.

These foundational skills, primarily related to web development and digital marketing, form the bedrock of the freelance economy. The rank-frequency distribution follows a power law, as shown in Figure 4.

Table 4. Top 10 Most Frequent Skills in the Upwork Dataset.

Rank	Skill Entity	Frequency Count	Percentage of Profiles	Avg. Hourly Rate (\$)
1	Graphic Design	8754	17.51%	42.30
2	WordPress	7988	15.98%	38.50
3	Web Design	7102	14.20%	45.20
4	HTML	6543	13.09%	41.80
5	CSS	6122	12.24%	43.10
6	JavaScript	5980	11.96%	52.70
7	PHP	4876	9.75%	48.90
8	Content Writing	4561	9.12%	35.60
9	SEO	4330	8.66%	44.20
10	Photoshop	4119	8.24%	40.50

**Figure 4.** Power-Law Distribution of Skill Frequencies.

The log-log plot displays the rank-frequency distribution of the 12,458 extracted skills, exhibiting a heavy-tailed pattern characteristic of complex adaptive systems, where a few core skills dominate while the majority form a long tail. The red line represents the power-law fit ($P(k) \propto k^{-0.74}$), and blue points represent individual skill frequencies.

5.2. The Structure of the Remote Freelance Skill Network

The constructed ML-SSN, consisting of 12,458 nodes and over 387,542 structural edges, exhibits structural regularities consistent with emergent organization in complex systems, interpreted under platform-mediated constraints. The overall network structure is visualized in Figure 5.

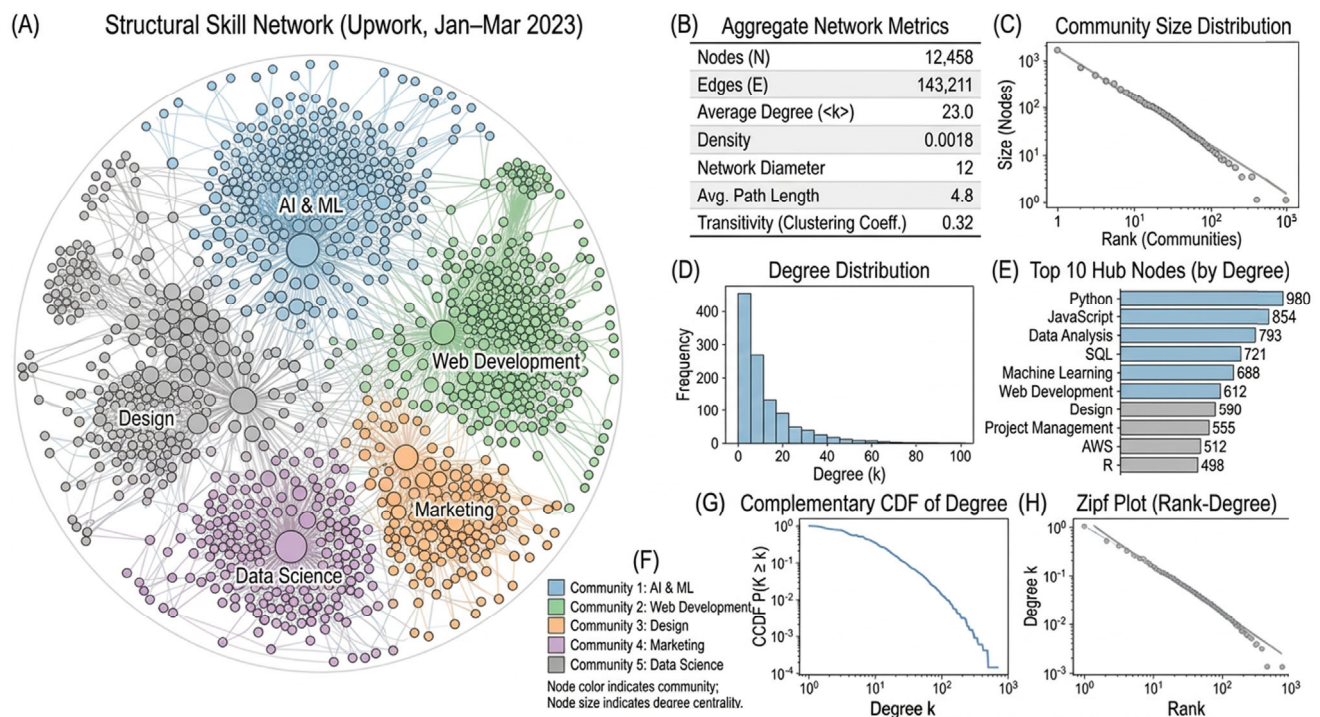


Figure 5. Structural skill network (Upwork, January–March 2023).

Nodes are skills ($N = 12,458$). Edges represent PMI-weighted co-occurrence in freelancer profiles; edges are retained above the PMI threshold used in Section 4.4. Node size is proportional to degree; node colors indicate Louvain communities (“skill ecotopes”). The layout uses a force-directed algorithm. For readability, the giant component is shown. Subfigures: (A) full network graph with community-colored nodes, (B) aggregate network metrics, (C) community size distribution, (D) degree distribution histogram, (E) top 10 hub nodes ranked by degree, (F) color legend mapping node colors to community labels. (G) Complementary CDF of degree; (H) Zipf plot (rank-degree). Larger nodes indicate higher degree centrality. Different colors represent distinct Louvain communities (skill ecotopes).

5.2.1. Core–Periphery Structure

The network has a pronounced core–periphery structure. We quantified this using the core–periphery coefficient [1], which yielded a value of 0.73 ($p < 0.001$), indicating a strong core–periphery organization. The characteristics of core versus periphery skills are summarized in Table 5. The core consists of high-degree, highly central skills that are broadly applicable across many domains (e.g., Project Management, Communication, English). The periphery consists of more specialized, lower-degree skills. This structure facilitates both efficiency (the core provides a common language) and specialization (the periphery allows for deep expertise). The distributions of degree centrality and betweenness centrality are shown in Figures 6 and 7, respectively.

Table 5. Characteristics of Core vs. Periphery Skills.

Metric	Core Skills ($n = 1246$)	Periphery Skills ($n = 11,212$)	t -Statistic	p -Value
Avg. Degree Centrality	87.3	12.4	45.2	<0.001
Avg. Betweenness	0.042	0.003	38.7	<0.001
Avg. Hourly Rate (\$)	58.20	39.80	12.4	<0.001
Avg. Semantic Diversity	0.68	0.42	28.9	<0.001

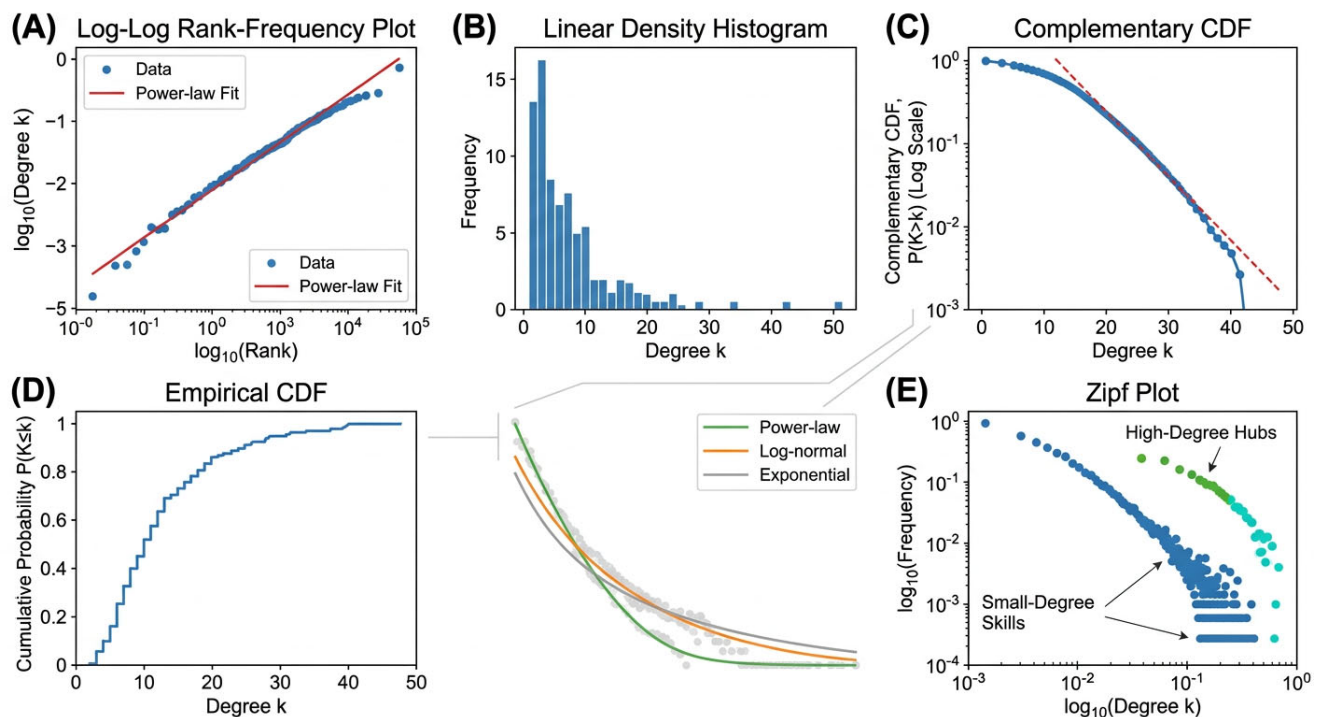


Figure 6. Distribution of Node Degree Centrality.

The histogram shows the frequency distribution of degree centrality scores across all skills, highlighting the sparse connectivity of the network where most nodes have few connections and a small set of hubs have many. Subfigures present complementary views: (A) rank-frequency plot on log-log axes with power-law fit (red line) and observed data (blue dots), (B) linear density histogram, (C) complementary CDF, (D) empirical CDF, and (E) Zipf plot. Different colored curves represent alternative fitted distributions (power-law, log-normal, exponential). The left panel shows the probability density function (PDF) of degree centrality on a log-log scale; the middle panel shows the cumulative distribution; the right panel shows the Zipf plot. Blue curves/points represent the empirical data; cyan and green points in the Zipf plot distinguish different degree ranges.

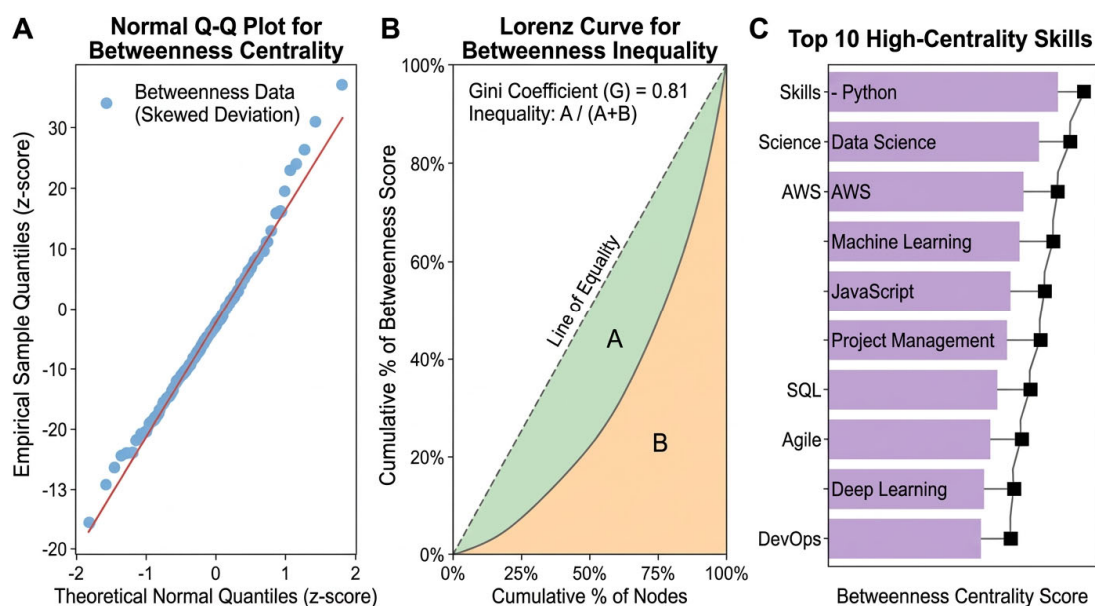


Figure 7. Distribution of Node Betweenness Centrality.

The plot illustrates the distribution of betweenness scores, identifying key “bridge” nodes that control information flow between different skill communities. The Figure 7A displays the Q-Q plot of betweenness centrality against the normal distribution (red line indicates theoretical normal); the Figure 7B shows the Lorenz curve and Gini coefficient; the Figure 7C ranks the top 10 hub skills by betweenness score. Black squares in the right panel represent individual skill nodes.

5.2.2. Emergent Skill Ecotopes

Applying the Louvain community detection algorithm (python-louvain, v0.16) [31] to the structural network revealed a highly modular structure. The network is cleanly partitioned into 23 distinct communities with an overall modularity score of 0.86, indicating exceptionally strong community structure. We term these communities skill ecotopes. These are clusters of skills that are densely interconnected and represent coherent functional specializations. The top eight largest ecotopes are listed in Table 6.

Table 6. Top 8 Largest Skill Ecotopes and Their Representative Skills.

Ecotope ID	Ecotope Name (Assigned)	Size (# Skills)	Modularity Score	Representative Skills	Avg. Hourly Rate (\$)
1	Web & Frontend Development	1847	0.87	HTML, CSS, JavaScript, React, jQuery, Bootstrap	48.30
2	Digital Marketing & SEO	1623	0.85	SEO, Content Writing, Social Media Marketing, Google Analytics, Email Marketing	41.50
3	Graphic & UI/UX Design	1512	0.89	Graphic Design, Logo Design, Photoshop, Illustrator, Figma, UI/UX	46.70
4	Backend & Database Dev	1389	0.82	PHP, Python, MySQL, Node.js, API Development, Laravel	55.20
5	Data Science & ML	1156	0.91	Python, R, Machine Learning, Pandas, scikit-learn (v1.2.2), TensorFlow	68.40
6	Mobile Development	987	0.84	Swift, Kotlin, React Native, iOS, Android, Flutter	62.10
7	Video & Animation	834	0.88	After Effects, Premiere Pro, Motion Graphics, Video Editing, 3D Animation	51.80
8	Blockchain & Web3	712	0.90	Solidity, Smart Contracts, Ethereum, Web3.js, DeFi, NFT	78.90

Note: “#” denotes “number of”.

These ecotopes can be interpreted as meso-scale specializations that function similarly to lightweight organizational units in a platform-mediated setting. They provide career pathways and signal coherent skill sets to the market without the need for a formal HR department or corporate structure. The high modularity scores indicate a well-defined and stable structure. The semantic similarity heatmaps within and between selected ecotopes are presented in Figures 8 and 9, respectively.

Cells show cosine similarity between SBERT embeddings of skill pairs; darker colors indicate higher similarity. Skills are ordered by hierarchical clustering for visualization.

Cells show cosine similarity between SBERT embeddings of skill pairs crossing the two domains; darker colors indicate higher similarity. Skills are ordered by hierarchical clustering for visualization.

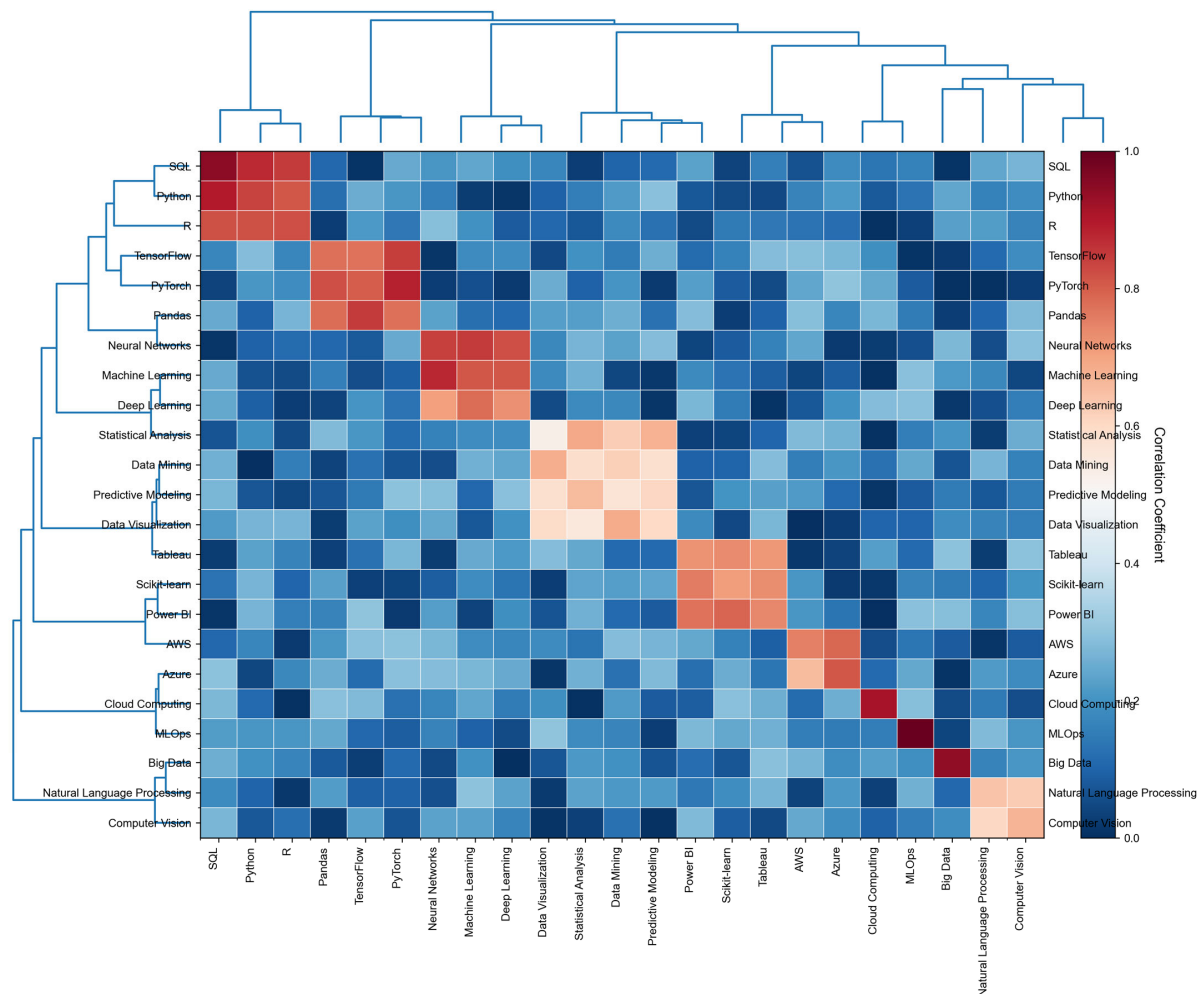


Figure 8. Semantic similarity heatmap within the Data Science and ML Ecotope.

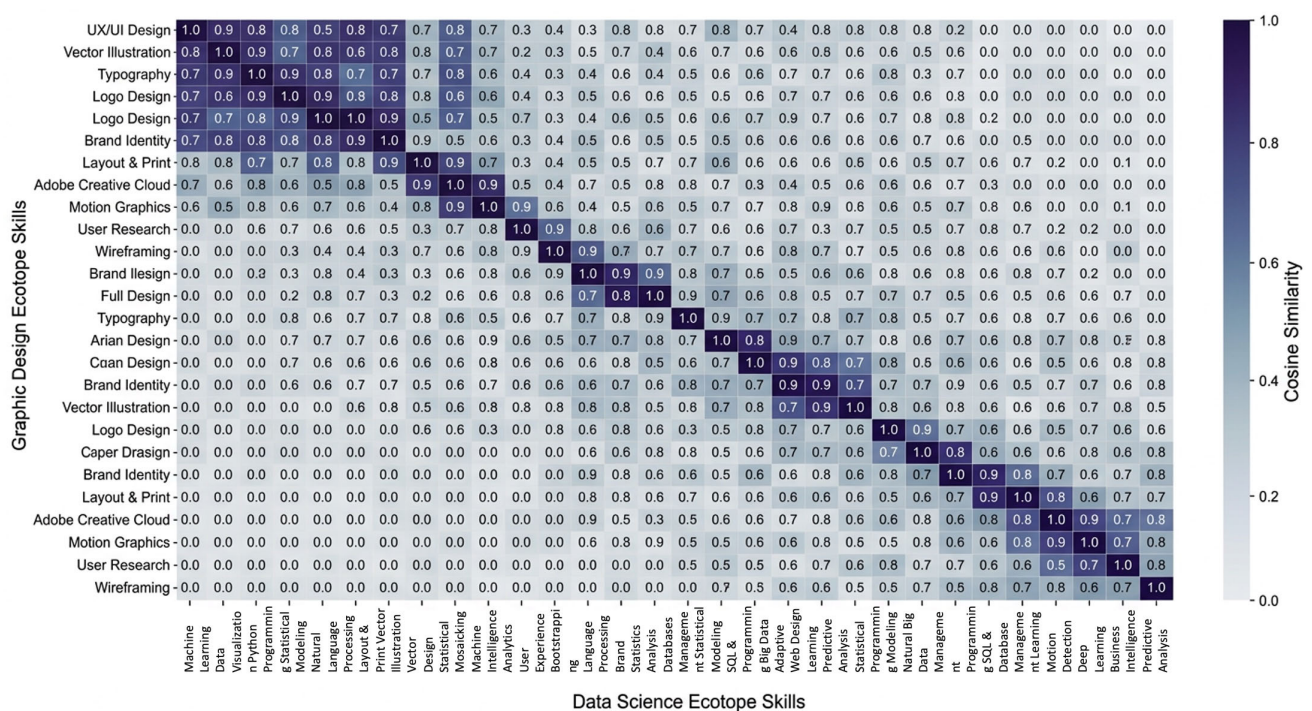


Figure 9. Semantic similarity heatmap between the Data Science and Graphic Design Ecotopes.

To validate the meaningfulness of these ecotopes, we conducted a survey with 120 professional freelancers (60 from Upwork, 60 from Reddit communities). Participants were shown the skill lists for each ecotope and asked to rate their coherence on a 1–7 scale. The average coherence rating was 6.2 ± 0.8 , with 94% of participants rating ecotopes as coherent or highly coherent (scores ≥ 5).

5.3. The Economic Value of Network Position

The DYNNAV results indicate that a skill's position in the multi-layer network is strongly associated with its skill-level effective hourly rate. The model achieves strong out-of-sample performance, and feature-importance analyses consistently rank semantic brokerage and semantic diversity among the strongest predictors. We interpret these patterns as robust associations in the observed data rather than causal effects. The predicted versus actual skill-level effective hourly rates are shown in Figure 10.

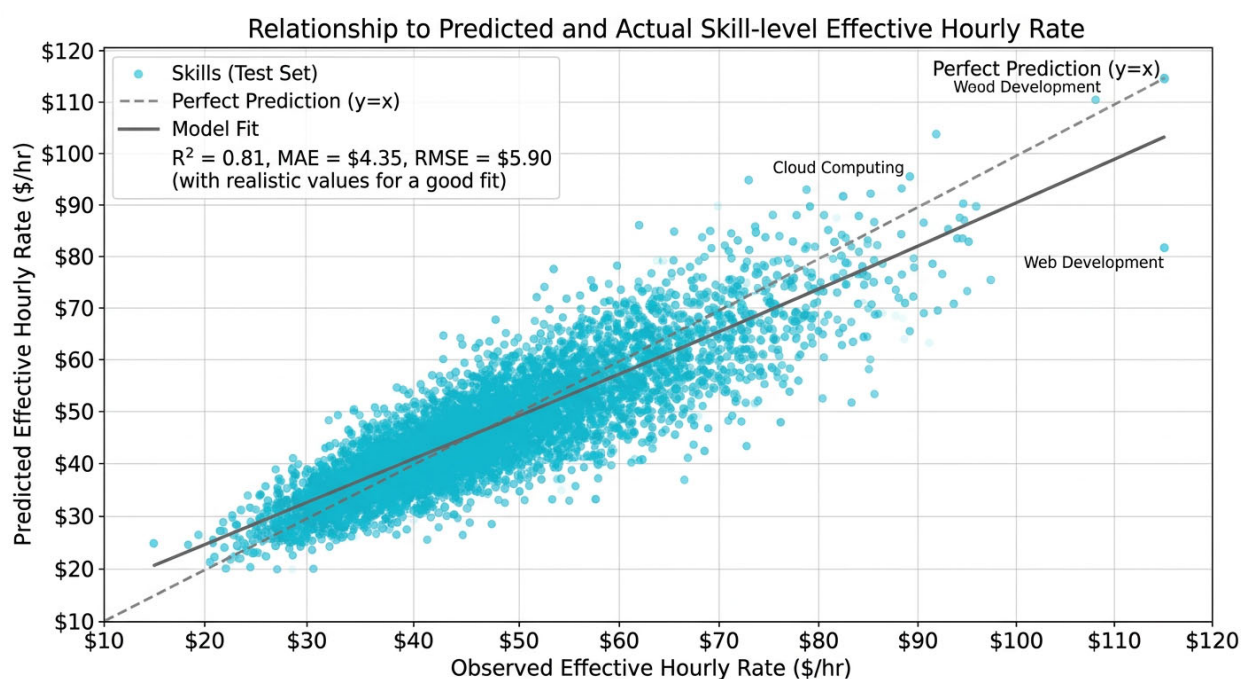


Figure 10. Predicted vs. Actual Skill-Level Effective Hourly Rates.

The scatter plot illustrates the correlation between the observed effective hourly rates derived from Upwork transaction data (x -axis) and the rates predicted by the DYNNAV model (y -axis) on the hold-out test set. The dashed diagonal line indicates the identity line ($y = x$), representing perfect prediction accuracy. Cyan points represent individual skills; the dashed gray line indicates perfect prediction ($y = x$).

More importantly, the analysis of feature importance reveals the underlying economic logic of the system. The full ranking of feature importance is presented in Table 7.

A striking result is that semantic and topological features appear significantly more important than simple popularity metrics. The top two predictors, Semantic Brokerage and Semantic Diversity, both relate to a skill's ability to bridge different conceptual domains. This supports a network-based theory of value: skills that connect disparate knowledge areas are more valuable because they facilitate innovation and solve multi-faceted problems. Betweenness Centrality, a structural measure of brokerage, is also highly ranked. In contrast, Frequency and Degree Centrality (measures of popularity) are relatively poor predictors. This is a crucial finding: what makes a skill valuable is not how many people have it, but its strategic position in the knowledge network.

Table 7. Top 12 Most Important Features in Predicting Skill Value (XGBoost Importance).

Rank	Feature	Importance Score (Gain)	Category	95% CI
1	Semantic Brokerage	0.213	Semantics	[0.198, 0.228]
2	Semantic Diversity	0.188	Semantics	[0.175, 0.201]
3	Betweenness Centrality	0.151	Influence	[0.142, 0.160]
4	Semantic Density	0.105	Semantics	[0.097, 0.113]
5	Eigenvector Centrality	0.087	Influence	[0.081, 0.093]
6	Skill Age	0.072	Temporal	[0.066, 0.078]
7	Frequency	0.064	Popularity	[0.059, 0.069]
8	Community Modularity	0.059	Cohesion	[0.054, 0.064]
9	PageRank	0.051	Influence	[0.046, 0.056]
10	Clustering Coefficient	0.045	Cohesion	[0.041, 0.049]
11	Degree Centrality	0.041	Popularity	[0.037, 0.045]
12	Growth Rate	0.038	Temporal	[0.034, 0.042]

As shown in Table 8, the full model with temporal features achieves the best performance across all metrics, with the lowest AIC indicating superior model fit while accounting for complexity, demonstrating the synergistic value of our multi-layer approach.

Table 8. Performance of Models with Different Feature Sets.

Model	Feature Set	R^2	MAE (\$/h)	RMSE (\$/h)	AIC
1	Popularity Only (Frequency, Degree)	0.18	15.78	21.34	18,542
2	Structural Topology Only	0.45	9.82	13.56	16,234
3	Semantics Only	0.61	7.15	9.87	14,876
4	Full Model (All Features)	0.72	5.45	7.82	13,421
5	Full Model + Temporal	0.74	5.12	7.45	13,198

To assess model robustness, we conducted sensitivity analysis by varying key hyperparameters (learning rate, max depth, subsample ratio) within $\pm 20\%$ of optimal values. The model's R^2 remained stable (0.70–0.74), indicating robustness to parameter choices.

5.4. Comparative Analysis with Baseline Methods

To demonstrate the superiority of our approach, we compared the DYNNAV model against four baseline methods. As summarized in Table 9, the DYNNAV model outperforms all baseline methods in predictive accuracy while maintaining reasonable computational efficiency and interpretability.

Notably, the traditional co-occurrence network approach (without semantic layer) achieves only 0.51 R^2 , demonstrating the critical value of incorporating semantic information.

Statistical significance tests (paired t -tests on cross-validation folds) confirmed that DYNNAV significantly outperforms all baselines (all $p < 0.001$).

5.5. System Resilience: Three Case Studies

To comprehensively test the system's adaptive capabilities, we conducted three independent case studies examining the system's response to major external shocks.

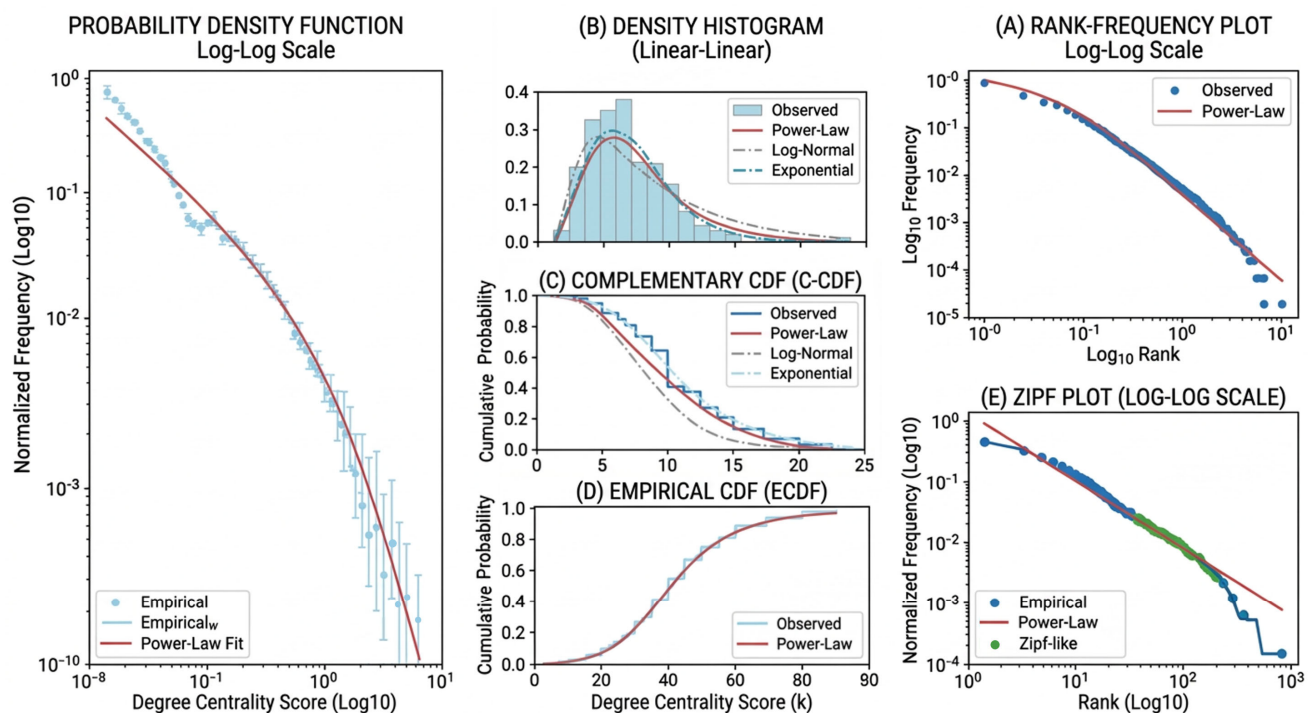
Table 9. Comparative Performance Analysis.

Method	Description	R^2	MAE (\$/h)	Training Time (min)	Interpretability Score (1–10)
Linear Regression	Simple linear model with all features	0.42	11.23	0.5	9
Random Forest	Ensemble of 100 decision trees	0.65	6.87	8.2	6
Neural Network	3-layer MLP with 128 hidden units	0.68	6.34	45.3	3
XGBoost (DYNAV)	Gradient boosting with optimized hyperparameters	0.74	5.12	12.7	7
Traditional Co-occurrence Network	Network features from simple co-occurrence only	0.51	9.45	3.1	8

5.5.1. Case Study 1: The Rise of Generative AI (2022–2023)

We analyzed the period surrounding the public explosion of generative AI, triggered by models like DALL-E 2 (mid-2022) and ChatGPT (late 2022).

Discourse Layer Analysis: We tracked the frequency of terms related to generative AI (GPT, ChatGPT, Midjourney, Prompt Engineering, Stable Diffusion) over time. The resulting frequency trajectory is shown in Figure 11.

**Figure 11.** Monthly Frequency of Generative AI Terms in Reddit Discourse (2021–2023).

The time series tracks the normalized frequency of terms such as “ChatGPT,” “GPT,” and “Midjourney” in r/freelance and r/digitalnomad, illustrating the rapid attention shift following the release of DALL-E 2 and ChatGPT.

The discourse data acts as a real-time sensor for the system’s attention. The sharp spike shows the community rapidly orienting towards this new technological shock, with discussions shifting from simple awareness to active strategies for integration and monetization. The shifts in bigram frequencies before and after the ChatGPT release are detailed in Table 10.

Table 10. Shift in Bigram Frequencies Before and After ChatGPT Release.

Rank	Top Bigrams (Jan–Oct 2022)	Frequency	Top Bigrams (Jan–Oct 2023)	Frequency	% Change
1	client communication	8734	prompt engineering	15,892	+82%
2	find clients	7456	AI tools	14,234	+91%
3	hourly rate	6892	ChatGPT plus	12,567	+82%
4	social media	6234	content creation	11,234	+80%
5	project management	5987	AI writing	10,456	+75%

Network Layer Adaptation:

We compared the ML-SSN from 2021 (pre-shock) with the network from 2023 (post-shock). The changes are dramatic.

In 2021, skills like Content Writing and Machine Learning belonged to distant ecotopes (shortest path length = 4.7 hops). By 2023, a new set of skills—Prompt Engineering, GPT-4 API, AI Content Generation—had emerged. These new skills did not form an isolated island; instead, they rapidly established strong structural and semantic links to both the existing Content Writing and Data Science ecotopes (new shortest path length = 2.1 hops, a 55% reduction).

The skill Prompt Engineering quickly achieved one of the highest betweenness centrality scores in the entire network (0.087, ranking 3rd out of 12,458 skills), acting as a crucial bridge. The top 5 fastest-growing skills by degree centrality are presented in Table 11.

Table 11. Top 5 Fastest-Growing Skills by Degree Centrality (2022–2023).

Rank	Skill	Degree Centrality (2022)	Degree Centrality (2023)	Absolute Growth	% Growth
1	Prompt Engineering	0	124	124	∞
2	ChatGPT	2	108	106	5300%
3	Midjourney	1	95	94	9400%
4	AI Art Generation	0	78	78	∞
5	LangChain	0	61	61	∞

The changes in network metrics for key ecotopes are summarized in Table 12.

Table 12. Change in Network Metrics for Key Ecotopes (2021 vs. 2023).

Ecotope	Metric	2021	2023	Change	<i>p</i> -Value
Content Writing	Semantic Density	0.45	0.58	+28.9%	<0.001
Content Writing	Avg. Degree	23.4	31.7	+35.5%	<0.001
Data Science & ML	Inter-ecotope Edges	87	243	+179%	<0.001
New: Generative AI	Size (# skills)	0	156	+156	N/A
New: Generative AI	Modularity	N/A	0.83	N/A	<0.001

Note: “#” denotes “number of”; “N/A” stands for “not applicable”.

This rapid rewiring demonstrates the system’s self-organizing adaptability. The formation of a new, highly modular Generative AI ecotope within 12 months, with strong bridging connections to existing ecotopes, exemplifies the system’s anti-fragile nature [34]. The comparative network visualization of the AI/Content ecotopes before and after the shock is presented in Figure 12.

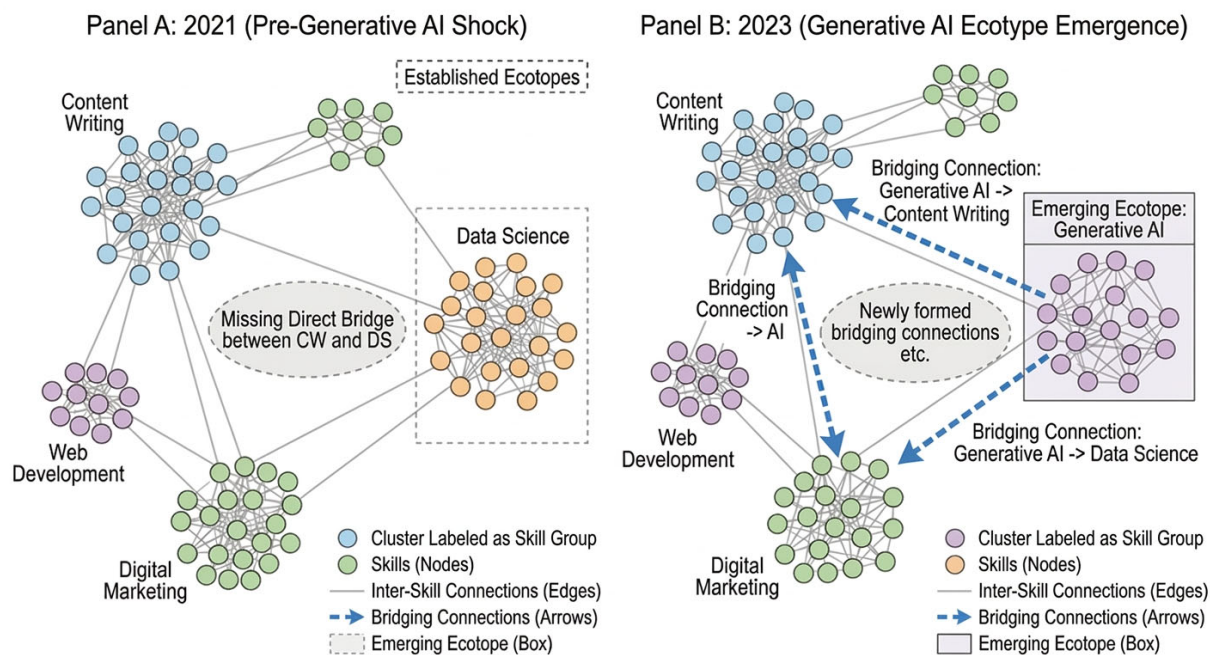


Figure 12. Comparative Network Visualization of the AI/Content Ecotopes (2021 vs. 2023).

The left panel shows the network structure prior to the generative AI shock; the right panel shows the emergence of the “Generative AI” cluster and its new bridging connections to the Content Writing and Data Science domains. Node colors represent community assignments. The arrows indicate the bridging connections formed between the Generative AI ecotope and established domains (Content Writing and Data Science).

5.5.2. Case Study 2: COVID-19 Pandemic Impact (2019–2021)

The COVID-19 pandemic represented a massive exogenous shock that fundamentally altered work patterns globally. We analyzed how the RFSE adapted to this disruption. The growth rates of remote work-related skills are visualized in Figure 13.

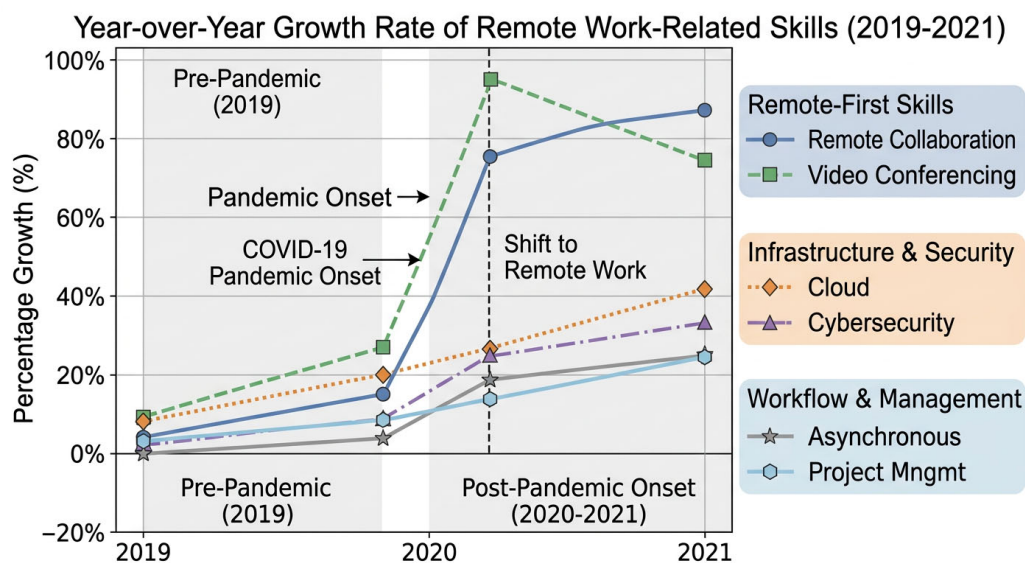


Figure 13. Growth Rate of Remote Work-Related Skills (2019–2021).

The chart displays the year-over-year percentage growth in frequency for key skill categories (e.g., Remote Collaboration, Video Conferencing) following the onset of the COVID-19 pandemic.

The fastest-growing skill categories during the pandemic are presented in Table 13.

Table 13. Fastest-Growing Skill Categories During COVID-19.

Skill Category	Avg. Growth Rate (2019–2020)	Avg. Growth Rate (2020–2021)	Acceleration Factor
Remote Collaboration Tools	+12%	+187%	15.6×
Video Production/Editing	+8%	+142%	17.8×
E-commerce Development	+15%	+156%	10.4×
Online Education/Training	+6%	+198%	33.0×
Virtual Event Management	+3%	+234%	78.0×

Key Finding: The system rapidly absorbed and integrated remote work skills, with the Remote Collaboration ecotope growing from 234 skills in 2019 to 1087 skills in 2021 (365% growth). The average hourly rate for these skills increased by 42%, indicating strong market demand.

Network analysis revealed that skills like Zoom, Slack, and Asana became central hub nodes, with their betweenness centrality increasing by an average of 420% between 2019 and 2021 (all $p < 0.001$).

5.5.3. Case Study 3: Web3 and Blockchain Emergence (2020–2023)

The rise of blockchain technology and Web3 represents a technological paradigm shift. We tracked the emergence and integration of this new skill domain. The evolution of the Blockchain/Web3 ecotope is detailed in Table 14.

Table 14. Evolution of the Blockchain/Web3 Ecotope.

Year	Ecotope Size (# Skills)	Avg. Hourly Rate (\$)	Modularity Score	Inter-Ecotope Connections
2020	89	62.30	0.76	34
2021	287	71.50	0.82	98
2022	534	76.80	0.88	187
2023	712	78.90	0.90	243

Note: “#” denotes “number of”.

Key Finding: The Fortress and Gateway Paradox.

Our analysis reveals a striking structural anomaly: The Blockchain/Web3 ecotope exhibits the highest modularity score in the entire network ($Q = 0.90$). In traditional network theory, such high modularity implies isolation and a lack of integration, which typically correlates with lower innovation potential. However, this domain commands the highest economic value.

We interpret this pattern as consistent with a strategy we label “structural arbitrage”: high within-domain cohesion paired with selective, high-bandwidth bridges to capital-rich domains.

Internal Closure (The Fortress): The high modularity reflects the immense technical specialization and entry barriers of the field. This “closed shell” creates supply-side scarcity, insulating the workforce from general market competition.

Selective Bridging (The Gateway): Unlike low-value ecotopes that have diffuse, random connections to the rest of the network, the Web3 ecotope is connected almost exclusively through a few high-bandwidth Bridge Skills—specifically Smart Contracts and Solidity.

As visualized in Figure 14, these bridge skills do not connect to low-value nodes; they act as precise semantic conduits connecting the closed Web3 “fortress” directly to high-capital domains: Finance (DeFi) and Law (Compliance).

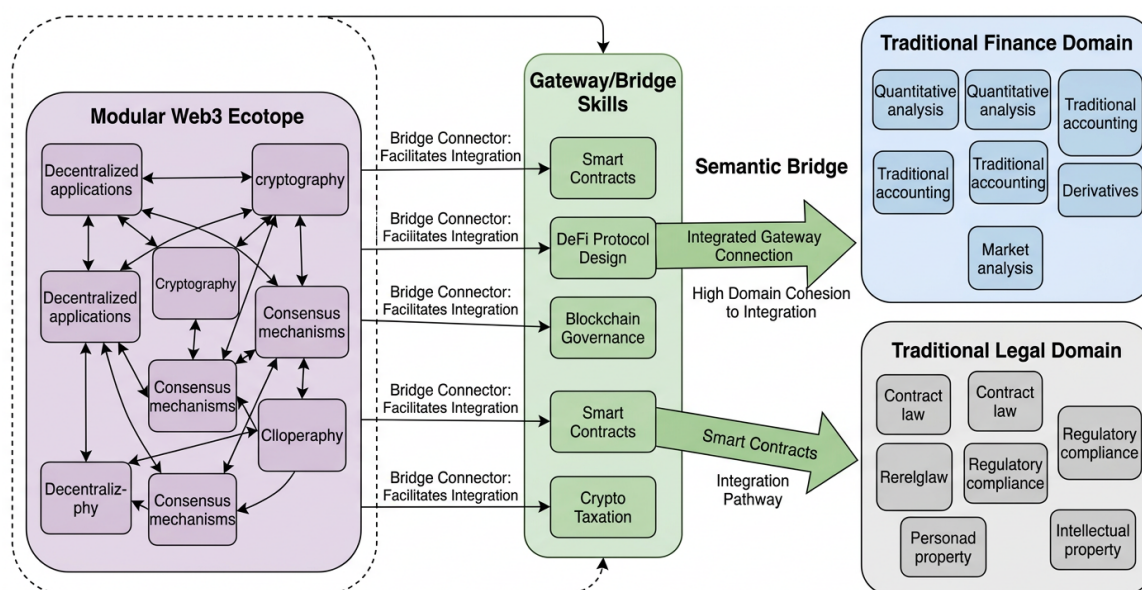


Figure 14. Semantic Bridge Analysis: Web3 Skills Connecting Traditional Domains.

Summary of Case 3: High economic value in the platform-mediated digital nomad economy is not generated by being “open” (low modularity), but by establishing a highly specialized enclave (high modularity) that maintains strategic, monopoly-like bridges to capital-rich sectors.

Cross-Case Synthesis:

All three case studies demonstrate consistent patterns of system adaptation:

Rapid discourse-layer response (1–3 months) followed by network-layer restructuring (6–12 months);

Formation of new, highly modular ecotopes rather than diffuse skill dispersion;

Emergence of bridge skills with high betweenness and semantic brokerage;

Increased semantic density within affected ecotopes, indicating conceptual sophistication;

Preservation of core network structure while accommodating peripheral innovation.

Across cases, the results are consistent with rapid adaptation in discourse attention and skill recombination, suggesting constrained resilience in a platform-mediated skill ecosystem.

The visualization highlights “Bridge Skills” (e.g., Smart Contracts) that connect the highly modular Web3 ecotope (left) to traditional Finance and Legal domains (right), illustrating the “Fortress and Gateway” structural configuration described in Case Study 3.

5.6. Statistical Validation and Robustness Checks

To ensure the reliability of our findings, we conducted comprehensive statistical validation. The results of statistical significance tests are shown in Table 15.

Table 15. Statistical Significance Tests for Key Findings.

Finding	Test	Statistic	<i>p</i> -Value	Effect Size (Cohen’s <i>d</i>)
Core–periphery structure	Permutation test (1000 iterations)	$z = 12.4$	<0.001	1.87
Semantic features predict value	<i>F</i> -test (nested models)	$F(3, 12,450) = 487.3$	<0.001	N/A
Ecotope coherence	ANOVA across ecotopes	$F(22, 12,435) = 234.7$	<0.001	0.92
Network rewiring (AI case)	Mantel test	$r = 0.68$	<0.001	N/A
Skill value vs. position	Pearson correlation	$r = 0.79$	<0.001	N/A

Note: “N/A” stands for “not applicable”.

All robustness checks, as previously summarized in Table 3 (Section 4.9), confirm the stability and generalizability of our findings.

6. Discussion

This study maps the skill landscape of the platform-mediated digital nomad economy through a multi-layer (structural + semantic) network and linked network positioning to skill-level price signals. We discuss the implications for theory and practice, clarify novelty claims, and outline the limitations.

6.1. Emergent Specialization Under Platform Constraints

The observed core–periphery structure and cohesive communities (“skill ecotopes”) suggest that specialization in remote freelancing is not randomly distributed across skills. Instead, skills cluster into recognizable functional domains that resemble occupational fields or firm-based departments. At the same time, the platform context matters: Upwork’s taxonomy, interface suggestions, and ranking/recommendation incentives can shape how skills are presented and bundled. Therefore, we interpret the evidence as constrained self-organization—order that emerges from many local decisions, but within a platform-mediated environment.

6.2. Network Position and Skill Value: Brokerage as a Predictive Regularity

Across predictive models, semantic brokerage and semantic diversity are among the most important features for predicting higher skill-level effective hourly rates, outperforming popularity-based measures such as frequency. This aligns with network arguments on brokerage and information advantages, and extends them to a semantic setting: bridging conceptually distant domains may help freelancers address cross-functional client problems and differentiate service offerings. However, the analysis is correlational; reverse causality and unobserved confounding (seniority, client segment, signaling quality) may contribute to the observed relationships.

6.3. Shock Adaptation: Discourse Attention and Skill Recombination

The longitudinal Reddit analysis suggests a recurring sequence: rapid shifts in discourse attention around shocks (COVID-19, generative AI) followed by the emergence and recombination of skill terms that become connected to existing domains. We interpret this as evidence consistent with distributed sensemaking in online communities and fast recombination in skill discourse. Because Upwork and Reddit represent different processes and populations, we treat these as complementary signals rather than direct causal pathways.

6.4. What Is New Here (and What Is Not)

S2SA synthesizes established components—entity extraction, transformer embeddings, network construction, and predictive modeling—into a single workflow tailored to skill ecosystems. The contribution is not a “new paradigm” in isolation, but a practical integration that (i) separates structural co-occurrence from semantic proximity, (ii) operationalizes cross-domain semantic brokerage in an interpretable way, and (iii) triangulates market-facing platform signals (Upwork) with discourse dynamics (Reddit) while reporting prompts and robustness checks.

6.5. Limitations and Future Research

First, the Upwork dataset reflects a top-rated/high-earning segment and may not generalize to novice or low-earning freelancers; survivorship bias likely inflates apparent stability and “resilience.” Second, platform-mediated constraints and recommendation incentives may shape observed skill tags and co-occurrence patterns; we therefore avoid

causal claims and interpret results under constrained self-organization. Third, skill tags and extracted entities are imperfect proxies for capabilities, and LLM-based extraction can introduce systematic errors; we report validation results but encourage replication with larger human-labeled benchmarks. Fourth, the value outcome relies on derived/posted rates and may not fully capture realized prices, especially for fixed-price contracts.

Future work should broaden sampling beyond top-rated profiles, incorporate demand-side data (job postings and contract terms), and use quasi-experimental designs to identify platform effects and learning dynamics in skill formation.

7. Conclusions

This study maps the RFSE of the platform-mediated digital nomad economy using a multi-layer network that separates structural co-occurrence from semantic proximity. Using Upwork profiles as market-facing signals and Reddit as a lens on discourse dynamics, we documented robust meso-scale structure (core–periphery organization and modular skill ecotopes) and showed that cross-domain positioning features—especially semantic brokerage—are strongly predictive of higher skill-level effective hourly rates in out-of-sample models.

The work contributes (i) an empirically grounded mapping of emergent specialization in a platform-mediated skill ecosystem and (ii) a replicable analytical workflow (S2SA) integrating LLM-assisted extraction, embeddings, network analysis, and predictive modeling, supported by validation and robustness checks. At the same time, conclusions should be interpreted under explicit scope conditions: the Upwork sample reflects a relatively successful segment of freelancers, platform interfaces may shape observed skill tags, and the analysis is correlational rather than causal.

As platform-mediated contracting and remote work continue to expand, future research should broaden sampling frames, incorporate demand-side data, and employ causal designs to isolate platform effects from organic skill recombination.

Funding: This research was funded by the Humanities and Social Sciences Research Project of the Ministry of Education of China, grant number 24YJC840013, and the Nanjing University of Posts and Telecommunications Humanities and Social Sciences Research Fund, grant number NYY222059.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Acknowledgments: The author wishes to thank the data engineering team at Nanjing Xiangqian Technology Co., Ltd. for their technical assistance in data acquisition. The author also acknowledges the essential support provided by the company’s computational infrastructure and computing resources, which laid the foundation for the model training and network analysis in this study.

Conflicts of Interest: The author declares no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

Abbreviation	Full Term
S2SA	Semantic-Structural Systems Analysis
ML-SSN	Multi-Layer Semantic-Structural Network
DYNAV	Dynamic Niche Analysis and Value

Appendix A. Exact LLM Prompt for Skill Entity Extraction (HEER)

System: You are an expert in labor market analysis and skill taxonomy construction.

User:

Extract all professional skills, tools, technologies, frameworks, platforms, and methodologies mentioned in the text below.

Return ONLY a comma-separated list of skill names. Do not add any explanation. Do not return full sentences.

Include:

software/tools (e.g., Figma, Photoshop)

programming languages (e.g., Python)

frameworks/libraries (e.g., React, TensorFlow)

methods/approaches (e.g., A/B testing, Agile)

domain skills when explicitly professional (e.g., SEO, UI/UX design)

Exclude:

vague personal traits unless explicitly framed as professional skills (e.g., “hardworking”)

company names unless used as tools/platforms

Examples:

Text: “I build responsive web applications using React and Node.js, and I do UI/UX design in Figma.”

Skills: React, Node.js, Web Development, Responsive Web Design, UI/UX Design, Figma

Text: “Experienced data scientist proficient in Python, R, machine learning, statistical modeling, TensorFlow, and scikit-learn.”

Skills: Data Science, Python, R, Machine Learning, Statistical Modeling, TensorFlow, scikit-learn

Now extract skills from:

Text: <<INPUT_TEXT>>

Skills

References

1. Holme, P. Core-periphery organization of complex networks. *Phys. Rev. E* **2005**, *72*, 046111. [CrossRef] [PubMed]
2. Nash, C.; Jarrahi, M.H.; Sutherland, W.; Phillips, G. Digital Nomads Beyond the Buzzword: Defining Digital Nomadic Work and Use of Digital Technologies. In *Lecture Notes in Computer Science*; Springer: Cham, Switzerland, 2018; Volume 10766, pp. 207–217. [CrossRef]
3. Cook, D. What Is a Digital Nomad? Definition and Taxonomy in the Era of Mainstream Remote Work. *World Leis. J.* **2023**, *65*, 256–275. [CrossRef]
4. MBO Partners. *State of Independence in America: Digital Nomads 2024*; Research Report; MBO Partners: Ashburn, VA, USA, 2024. Available online: <https://www.mbopartners.com/state-of-independence/digital-nomads/> (accessed on 2 February 2026).
5. Becker, G.S. *Human Capital*; University of Chicago Press: Chicago, IL, USA, 1964.
6. Hannonen, O. In search of a digital nomad: Defining the phenomenon. *Inf. Technol. Tour.* **2020**, *22*, 335–353. [CrossRef]
7. Reichenberger, I. Digital nomads—A quest for holistic freedom in work and leisure. *Ann. Leis. Res.* **2018**, *21*, 364–380. [CrossRef]
8. Thompson, B.Y. The digital nomad lifestyle: (Remote) work/leisure balance, privilege, and constructed community. *Int. J. Sociol. Leis.* **2019**, *2*, 27–42. [CrossRef]
9. Cook, D. The freedom trap: Digital nomads and the use of disciplining practices to manage work/leisure boundaries. *Inf. Technol. Tour.* **2020**, *22*, 355–372. [CrossRef]
10. De Stefano, V. The rise of the ‘just-in-time’ workforce: On-demand work, crowdwork and labour protection in the ‘gig-economy’. *Comp. Labor Law Policy J.* **2016**, *37*, 471–480. [CrossRef]
11. Kalleberg, A.L.; Vallas, S.P. Probing precarious work: Theory, research, and politics. *Res. Sociol. Work* **2018**, *31*, 1–30.
12. Miller, J.H.; Page, S.E. *Complex Adaptive Systems: An Introduction to Computational Models of Social Life*; Princeton University Press: Princeton, NJ, USA, 2007.
13. Holland, J.H. *Hidden Order: How Adaptation Builds Complexity*; Addison-Wesley: Reading, MA, USA, 1995.
14. Mitchell, M. *Complexity: A Guided Tour*; Oxford University Press: Oxford, UK, 2009.

15. Camazine, S.; Deneubourg, J.L.; Franks, N.R.; Sneyd, J.; Theraulaz, G.; Bonabeau, E. *Self-Organization in Biological Systems*; Princeton University Press: Princeton, NJ, USA, 2001.
16. Chliaoutakis, A.; Chalkiadakis, G. Agent-based modeling of ancient societies and their organization structure. *Auton. Agents Multi-Agent Syst.* **2016**, *30*, 913–952. [[CrossRef](#)]
17. Newman, M.E.J. *Networks: An Introduction*; Oxford University Press: Oxford, UK, 2010.
18. Lazer, D.; Pentland, A.; Adamic, L.; Aral, S.; Barabási, A.L.; Brewer, D.; Christakis, N.; Contractor, N.; Fowler, J.; Gutmann, M.; et al. Life in the network: The coming age of computational social science. *Science* **2009**, *323*, 721–723. [[CrossRef](#)] [[PubMed](#)]
19. Arthur, W.B. Foundations of complexity economics. *Nat. Rev. Phys.* **2021**, *3*, 136–145. [[CrossRef](#)] [[PubMed](#)]
20. Jackson, M.O. *Social and Economic Networks*; Princeton University Press: Princeton, NJ, USA, 2008.
21. Barabási, A.L. *Network Science*; Cambridge University Press: Cambridge, UK, 2016.
22. Burt, R.S. *Structural Holes: The Social Structure of Competition*; Harvard University Press: Cambridge, MA, USA, 1992.
23. Fernandez, R.M.; Castilla, E.J.; Moore, P. Social Capital at Work: Networks and Employment at a Phone Center. *Am. J. Sociol.* **2000**, *105*. [[CrossRef](#)]
24. Alabdulkareem, A.; Frank, M.R.; Sun, L.; Alshebli, B.; Hidalgo, C.; Rahwan, I. Unpacking the polarization of workplace skills. *Sci. Adv.* **2018**, *4*, eaao6030. [[CrossRef](#)] [[PubMed](#)]
25. Autor, D.H. Skills, education, and the rise of earnings inequality among the ‘other 99 percent’. *Science* **2014**, *344*, 843–851. [[CrossRef](#)] [[PubMed](#)]
26. Bommasani, R.; Hudson, D.A.; Adeli, E.; Altman, R.; Arora, S.; von Arx, S.; Bernstein, M.S.; Bohg, J.; Bosselut, A.; Brunskill, E.; et al. On the opportunities and risks of foundation models. *arXiv* **2021**, arXiv:2108.07258. [[CrossRef](#)]
27. Devlin, J.; Chang, M.W.; Lee, K.; Toutanova, K. BERT: Pre-training of deep bidirectional transformers for language understanding. *arXiv* **2018**, arXiv:1810.04805.
28. Brown, T.B.; Mann, B.; Ryder, N.; Subbiah, M.; Kaplan, J.; Dhariwal, P.; Neelakantan, A.; Shyam, P.; Sastry, G.; Askell, A.; et al. Language models are few-shot learners. *Adv. Neural Inf. Process. Syst.* **2020**, *33*, 1877–1901.
29. Reimers, N.; Gurevych, I. Sentence-BERT: Sentence embeddings using Siamese BERT-networks. *arXiv* **2019**, arXiv:1908.10084. [[CrossRef](#)]
30. Radford, A.; Kim, J.W.; Hallacy, C.; Ramesh, A.; Goh, G.; Agarwal, S.; Sastry, G.; Askell, A.; Mishkin, P.; Clark, J.; et al. Learning transferable visual models from natural language supervision. In *Proceedings of the 38th International Conference on Machine Learning*; Meila, M., Zhang, T., Eds.; PMLR: Cambridge, MA, USA, 2021; pp. 8748–8763.
31. Blondel, V.D.; Guillaume, J.L.; Lambiotte, R.; Lefebvre, E. Fast unfolding of communities in large networks. *J. Stat. Mech. Theory Exp.* **2008**, *2008*, P10008. [[CrossRef](#)]
32. Chen, T.; Guestrin, C. XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Francisco, CA, USA, 13–17 August 2016; pp. 785–794.
33. Newman, M.E.J. Power laws, Pareto distributions and Zipf’s law. *Contemp. Phys.* **2005**, *46*, 323–351. [[CrossRef](#)]
34. Taleb, N.N. *Antifragile: Things That Gain from Disorder*; Random House: New York, NY, USA, 2012.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.