Computers, Materials & Continua

Tech Science Press

https://doi.org/10.32604/cmc.2026.086887

ARTICLE

Cross-Provider OAuth Capability Topology: A Structural Network Analysis of Modern Authorization Ecosystems

Maryam Almarwani*

Department of Cybersecurity, College of Computer Science and Engineering, Taibah University, Medina, Saudi Arabia

*Corresponding Author: Maryam Almarwani. Email: mmmarwani@taibahu.edu.sa

Received: 06 June 2026; Accepted: 12 August 2026; Published: 15 September 2026

ABSTRACT: OAuth authorization ecosystems contain a large and diverse collection of capabilities distributed across multiple cloud platforms. Although previous studies have investigated OAuth security, privacy, and authorization management, the structural organization of authorization capabilities across providers has received limited attention. This study presents a cross-provider structural analysis of OAuth capabilities from seven major authorization platforms. A unified capability dictionary is constructed by normalizing publicly documented OAuth scopes into a common semantic representation. The normalized capabilities are transformed into an undirected semantic topology in which nodes represent capabilities and edges represent deterministic semantic relationships. Standard network analysis and community detection techniques are then applied to examine connectivity patterns, central capabilities, structural communities, and cross-provider bridges. The resulting topology contains 1599 capabilities connected by 6049 unique semantic relationships, forming 525 semantic communities with a modularity value of 0.9045 and 2307 cross-provider bridge relationships. Robustness analyses using random, degree-preserving, provider-label-permuted, text-similarity, and alternative rule-based topologies show that the observed modular organization cannot be explained by graph size or degree distribution alone, while the absolute number of cross-provider bridges remains sensitive to the adopted semantic matching criterion. The proposed methodology provides a reproducible structural representation of OAuth authorization ecosystems and can support future studies integrating semantic topology with operational authorization data.

KEYWORDS: OAuth; OAuth scopes; authorization; network analysis; community detection; semantic topology; access control; cloud security

1 Introduction

OAuth has become a widely adopted authorization protocol for web and cloud applications [1]. It enables users to grant limited access to third-party services without sharing account credentials. Major platforms such as Google, Microsoft, GitHub, Slack, Dropbox, Discord, and Notion expose large collections of authorization scopes to support different application functions [2–4]. Each OAuth scope represents a specific authorization capability, providing access to user identity information, files, communication services, calendars, or organizational resources. As cloud ecosystems continue to expand, the number and diversity of available scopes also increase. This growth creates authorization environments containing many related capabilities distributed across multiple providers [5–7]. Existing studies have mainly investigated OAuth security, protocol implementation, privacy implications, and authorization management. Previous work analyzed implementation weaknesses, authorization misuse, privacy risks, and deployment characteristics across

Copyright © 2026 The Author. Published by Tech Science Press. This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

different platforms [1,3,7–10]. Other studies explored authorization structures and semantic management of scopes [5,11]. However, most existing work examined individual scopes or provider-specific authorization models rather than the structural organization of capabilities across multiple providers. Network analysis provides a natural framework for studying relationships between connected entities. Community detection and topology analysis have been successfully applied to software dependency networks, social networks, and other complex systems to identify structural groups and connectivity patterns [12–15]. These methods have received limited attention in OAuth authorization ecosystems, especially for analyzing semantic relationships between capabilities provided by different authorization platforms. This study constructs a unified cross-provider OAuth capability topology using publicly available scope definitions collected from multiple authorization providers. The resulting network represents semantic relationships between capabilities and supports analysis of communities, central capabilities, connectivity patterns, and cross-provider bridges. The analysis focuses on the structural organization of authorization capabilities rather than protocol implementation or application behavior. The study addresses the following research questions.

- RQ1: How are OAuth capabilities structurally organized across major authorization providers?
- RQ2: Which capability communities consistently emerge across heterogeneous OAuth ecosystems?
- RQ3: Which capabilities function as structural bridges between providers and authorization domains?
- RQ4: What structural characteristics describe concentration and fragmentation within modern OAuth authorization ecosystems?

Unlike previous studies that primarily investigate OAuth protocol security, deployment characteristics, or provider-specific scope management, this work addresses the structural organization of authorization capabilities across heterogeneous providers. Rather than analyzing protocol behavior or permission usage, the proposed framework constructs a unified semantic topology that enables reproducible graph-based analysis of capability communities, structural centrality, and cross-provider semantic bridges using publicly documented OAuth capabilities. The scientific novelty of this work lies in the integration of four complementary components within a single reproducible analytical framework: (i) a unified cross-provider OAuth capability dictionary, (ii) a deterministic semantic normalization process for heterogeneous authorization scopes, (iii) a graph-based structural representation enabling reproducible topology analysis, and (iv) a systematic characterization of cross-provider semantic bridges. Collectively, these components provide a structural perspective on OAuth authorization ecosystems that is not addressed by prior protocol-centric or provider-specific studies. The main contributions of this work are summarized as follows

- A reproducible semantic normalization framework that unifies heterogeneous OAuth capability definitions from seven major authorization providers into a common representation.
- A unified cross-provider semantic topology that enables provider-independent structural analysis of OAuth authorization ecosystems.
- A graph-based analytical framework that characterizes semantic communities, structurally central capabilities, and cross-provider semantic bridges.
- A reproducible structural representation that supports authorization auditing, semantic capability comparison, and future integration with operational authorization data.

The remainder of this paper is organized as follows. Section 2 reviews related work; Section 3 presents the methodology; Section 4 reports the results; Section 5 discusses the findings; Section 6 outlines the limitations; and Section 7 concludes the paper.

2 Related Work

2.1 OAuth Security and Authorization

OAuth has been widely studied as a standard authorization protocol for web and cloud applications. Previous studies examined protocol implementation, authorization flows, application impersonation, security weaknesses, security testing, and authorization compliance across OAuth ecosystems [2,3,6,16–18]. More recent work has further strengthened OAuth and OpenID Connect security through updated security recommendations and formal verification. RFC 9700 consolidates current best practices for OAuth 2.0 security, while AuthSaber automatically verifies OpenID Connect implementations against formal safety specifications and identifies previously unknown vulnerabilities in widely deployed libraries [19]. Collectively, these studies improved the understanding of OAuth security, implementation, privacy, and deployment characteristics [7,9], but they do not investigate the structural organization or cross-provider relationships of authorization capabilities.

2.2 OAuth Scope Management

OAuth scopes define the permissions granted to third-party applications. Previous studies explored authorization management, semantic organization, scope misuse, authorization assumptions, and permission interpretation [5,8,11,20], but generally treated scopes as independent authorization units or provider-specific permission models, leaving cross-provider semantic relationships and graph-based representations relatively unexplored. Recent studies have also investigated cloud identity and access management (IAM), permission engineering, least-privilege authorization, semantic access-control analysis, and permission recommendation for large-scale cloud environments [21,22]. These studies focus on permission assignment, policy optimization, and authorization enforcement rather than the structural organization of heterogeneous OAuth capability ecosystems, complementing rather than replacing graph-based semantic topology analysis across multiple authorization providers.

2.3 Network-Based Structural Analysis

Network analysis, including community detection and topology analysis, has been widely applied to software dependency networks, social networks, information systems, package ecosystems, and software dependency graphs to identify structural groups and connectivity patterns [12–15,23,24]. However, similar structural analysis has rarely been applied to OAuth authorization capabilities across multiple providers.

2.4 Research Gap

Although prior studies provide extensive analysis of OAuth security, privacy, implementation, authorization management, and network-based analysis, few combine these directions by representing OAuth capabilities as a unified cross-provider semantic topology. This study addresses that gap by constructing a semantic network from multiple authorization providers to analyze capability communities, central capabilities, connectivity patterns, and cross-provider bridges. Unlike existing studies on OAuth security, IAM policy analysis, or semantic permission engineering [19,22], this work focuses on constructing a reproducible structural representation of authorization capabilities spanning multiple providers. The contribution is therefore not a new authorization mechanism, but a graph-based analytical representation that supports capability comparison, structural auditing, and future integration with operational authorization evidence.

3 Methodology

This study follows a reproducible five-stage workflow for constructing and analyzing a cross-provider OAuth capability topology: capability collection, normalization, semantic topology construction, network and community analysis, and structural bridge analysis. Fig. 1 summarizes the methodology.

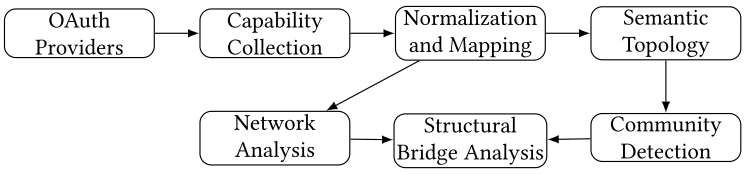

Figure 1: Five-stage workflow for constructing and analyzing the OAuth capability topology.

3.1 OAuth Capability Collection

OAuth capability definitions were collected from publicly available documentation for Google, Microsoft, GitHub, Slack, Dropbox, Discord, and Notion. Only publicly documented capabilities were included in the unified dictionary. Each capability was stored with its provider, name, resource category, access level, and source reference, and duplicate entries were removed during preprocessing to produce a unified capability dictionary.

3.2 Capability Normalization

Different providers use different naming conventions for authorization scopes. A normalization step was therefore applied before topology construction. Each capability was represented by five attributes: provider, capability name, resource category, access level, and sensitivity category. Resource names were converted to a consistent representation, and access levels were mapped into common categories including read, write or administrative access, identity-related access, and other capabilities. Normalization rules were applied consistently across all providers using predefined mapping criteria, with ambiguous cases manually reviewed against the official provider documentation before inclusion in the unified capability dictionary. The complete normalization rules, category definitions, representative and ambiguous mapping cases, and the complete normalization taxonomy are provided in the accompanying Supplementary Material to support reproducibility and external inspection. As a documentation-based consistency audit, a stratified random sample of 30 OAuth capabilities covering all seven authorization providers was rechecked against the corresponding official provider documentation for resource category, access level, and sensitivity category, with no inconsistencies identified. This audit evaluates consistency with the declared normalization criteria but does not constitute independent multi-annotator validation; no second annotator or inter-annotator agreement statistic was available in the present study. Table 1 summarizes the documentation-based consistency audit used to assess the semantic normalization process.

Table 1: Summary of the documentation-based normalization consistency audit.

Audit Item	Value
Audit strategy	Stratified random sampling
Sample size	30 capabilities
Authorization providers	7
Audit source	Official provider documentation
Verified attributes	Resource, Access, Sensitivity
Outcome	No inconsistencies identified

3.3 Semantic Topology Construction

The normalized capability dictionary was transformed into an undirected semantic network in which each node represents an OAuth capability and each edge represents a semantic relationship. The topology is intentionally constructed as an attribute-based semantic representation rather than an automatically inferred semantic graph, providing a reproducible structural abstraction based on explicitly defined semantic attributes. Edges were created using four deterministic rules.

1. Capabilities belonging to the same provider and resource category.
2. Capabilities sharing the same provider, sensitivity category, and access level.
3. Cross-provider capabilities with matching resource semantics.
4. Identity-sensitive capabilities belonging to the same provider.

Because Rule 3 directly determines cross-provider connectivity, we additionally evaluated its sensitivity using a stricter alternative formulation requiring exact normalized-resource agreement between capabilities from different providers. This alternative was used to assess whether the reported structural conclusions remain stable under a more restrictive cross-provider matching criterion. These deterministic rules generate a reproducible structural representation of authorization capabilities without application logs or deployment-specific information. Community detection is then applied to characterize the resulting structural organization rather than rediscover the predefined semantic attributes.

3.4 Network Analysis

Structural organization was characterized using degree (direct connectivity), weighted degree (connection strength), PageRank (influence), betweenness centrality (bridge roles), and closeness centrality (network proximity).

3.5 Community Detection

Community detection was performed using the Louvain algorithm because of its efficiency for modularity optimization in large sparse networks [24,25]. Detected communities were summarized by size, dominant provider, dominant sensitivity category, and average structural metrics.

3.6 Structural Bridge Analysis

The final stage identifies structural bridges between authorization providers. A bridge is defined as an edge connecting capabilities from different providers that share compatible semantic characteristics. These bridges indicate semantic compatibility rather than authorization equivalence, providing a structural view of heterogeneous OAuth ecosystems. The workflow is fully reproducible because every step is deterministic and based on publicly available capability definitions.

4 Experimental Results

This section presents the structural analysis of the constructed OAuth capability topology, including the capability dataset, network structure, central capabilities, community organization, and cross-provider bridge relationships.

4.1 Capability Dataset Overview

Table 2 summarizes the constructed topology and the distribution of normalized OAuth capabilities across seven authorization providers.

Table 2: Summary of the constructed OAuth capability topology and provider distribution.

(a) Topology Summary	
Metric	**Value**
OAuth capabilities	1599
Authorization providers	7
Topology nodes	1599
Semantic edges	6049
Connected components	514
Largest connected component	517
Detected communities	525
Cross-provider bridges	2307
Average clustering coefficient	0.5437
Modularity	0.9045
Density	0.0047
(b) Provider Distribution	
Provider	**Capabilities**
Microsoft	964
Google	467
Slack	72
GitHub	37
Discord	28
Dropbox	23
Notion	8

4.2 Provider Distribution

Fig. 2 shows that Microsoft contributes the largest number of documented capabilities, followed by Google, while the remaining providers contribute smaller but diverse capability sets. The potential influence of this imbalance is examined through the provider sensitivity analysis in Section 5.

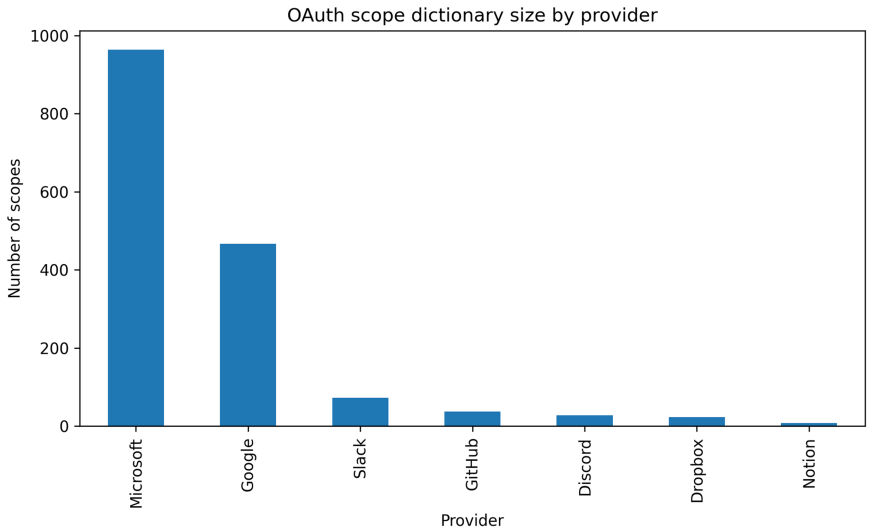

Figure 2: Distribution of OAuth capabilities across authorization providers.

4.3 Network Topology

Global network statistics indicate a sparse semantic topology with 1599 nodes, 6049 relationships, and a density of 0.0047. The topology contains 514 connected components, including one dominant component comprising 517 capabilities. The average clustering coefficient of 0.5437 indicates frequent local semantic grouping rather than a single homogeneous structure.

4.4 Central Capability Analysis

PageRank, weighted degree, betweenness, and closeness centrality were calculated to identify structurally important capabilities. Fig. 3 presents the highest ranked capabilities according to PageRank.

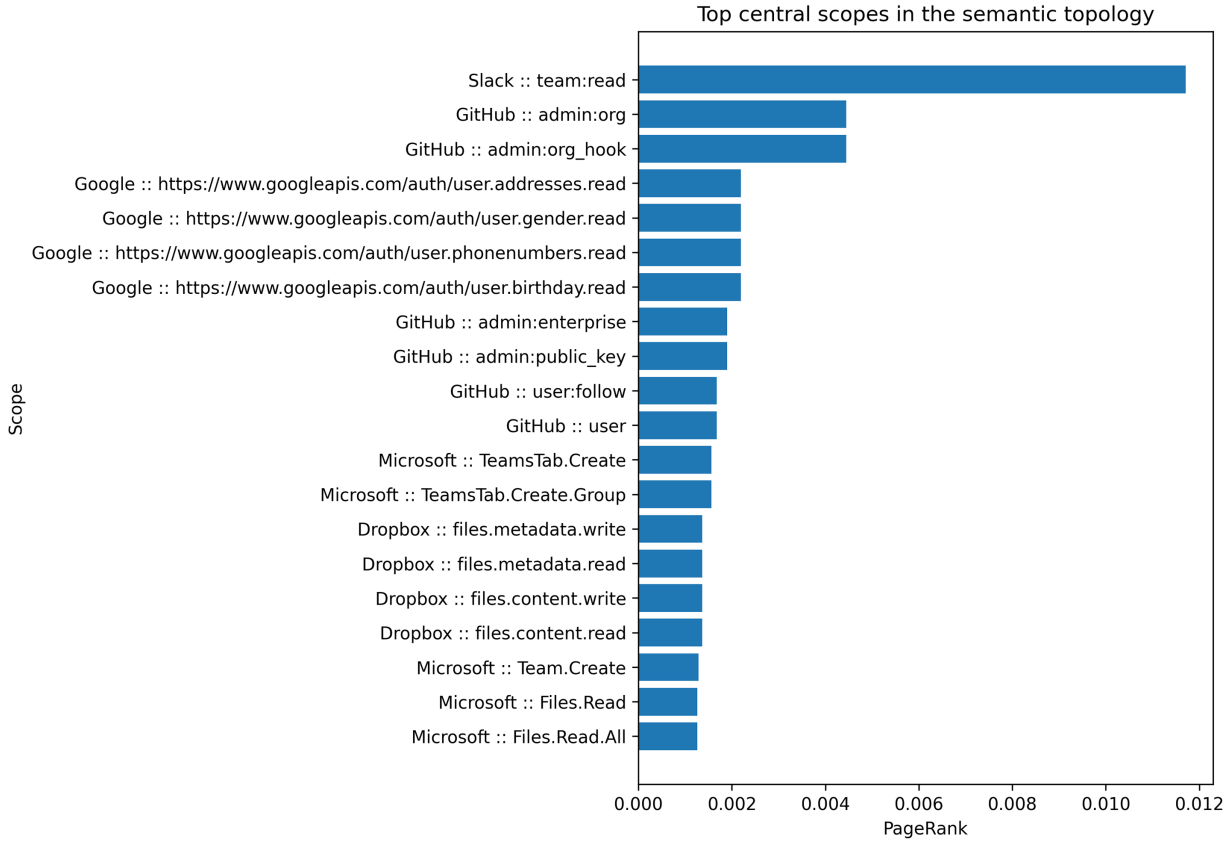

Figure 3: Top OAuth capabilities ranked by PageRank.

Organization, identity, and content capabilities dominate the highest PageRank positions, connecting multiple semantic regions.

4.5 Capability Communities

Community detection using the Louvain algorithm identified 525 communities with a modularity value of 0.9045. Fig. 4 presents the largest detected communities. Community sizes are highly uneven, with a few large communities and many smaller ones. Their semantic interpretation is provided in Section 5.

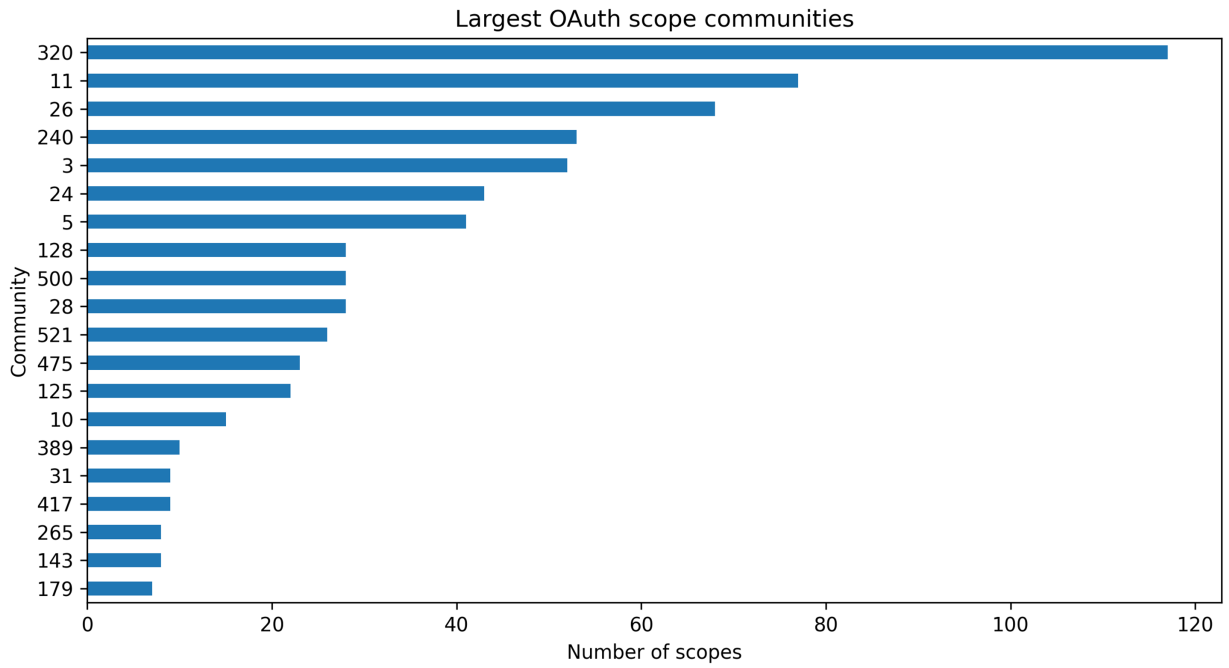

Figure 4: Largest capability communities identified in the OAuth topology.

4.6 Cross-Provider Capability Bridges

The constructed topology contains 2307 cross-provider bridge relationships under the declared semantic matching rules. These edges represent candidate structural correspondences between capabilities with compatible resource semantics rather than independently discovered authorization equivalences. Their sensitivity to alternative Rule 3 definitions is evaluated in Section 5.

5 Discussion

5.1 Structural Organization of OAuth Capabilities

The first research question investigates how OAuth capabilities are organized across authorization providers. The constructed topology is sparse, with a density of 0.0047, while maintaining an average clustering coefficient of 0.5437. These values indicate that capabilities are not uniformly connected. Instead, they form multiple local structures with stronger internal relationships. The network contains 514 connected components, including one dominant component comprising 517 capabilities. Under the declared semantic representation, this structure separates OAuth capabilities into multiple semantic regions rather than a single homogeneous authorization space. Fig. 5 illustrates the global semantic organization of the constructed topology. For visualization clarity, the eight largest detected communities are highlighted using distinct colors, while the remaining communities are shown in gray. The visualization demonstrates that OAuth capabilities are organized into several dense semantic regions connected through a limited number of bridge capabilities.

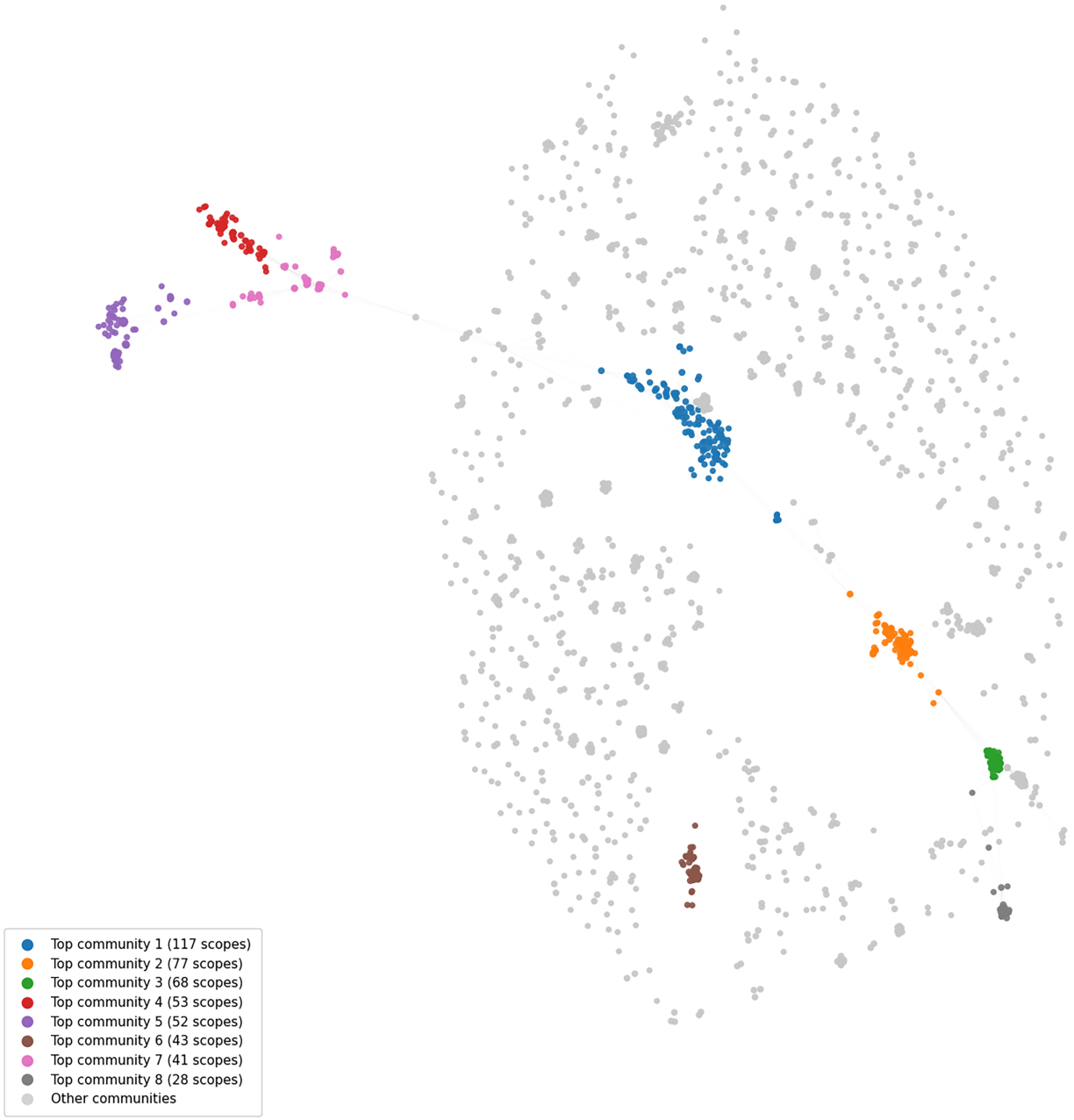

Figure 5: Force-directed visualization of the semantic OAuth capability topology. Colored nodes represent the eight largest communities, gray nodes denote the remaining communities, and node size is proportional to PageRank centrality.

To evaluate whether the reported structural observations could be explained by graph-construction choices, node-degree structure, or provider-scale effects, complementary robustness analyses were performed using random graphs, configuration-model and exact degree-preserving randomizations, a density-matched provider-label permutation, a text-similarity topology, provider exclusion, and edge-rule sensitivity. Each stochastic baseline was evaluated over 200 realizations using seeds 42–241, one per realization, and results are reported as mean ± sample standard deviation. Random graphs preserved the observed node and edge counts, configuration-model networks preserved the degree sequence in expectation, and

exact degree-preserving randomizations used $10|E|$ double-edge swaps per realization. For the provider-label test, labels were permuted while preserving provider counts, an alternative topology was reconstructed using exact normalized-resource agreement for cross-provider connectivity, and the resulting graph was density-matched to 6049 edges. The text-similarity topology used TF–IDF representations of capability names and available descriptions with unigram and bigram features, retaining the 6049 highest positive cosine-similarity edges. The results are summarized in Table 3.

Table 3: Robustness analyses including null-model comparisons, alternative graph constructions, provider sensitivity, and edge-rule sensitivity.

(a) Provider Sensitivity Analysis

Metric	Full	Without Microsoft
Nodes	1599	635
Edges	6049	2246
Communities	525	205
Modularity	0.9045	0.8956
Cross-provider bridges	2307	248

(b) Edge-Rule Ablation

Configuration	Communities	Modularity	Bridges
Full topology	525	0.9045	2307
Without Rule 1	1071	0.7497	2307
Without Rule 3	657	0.9563	0
Strict Rule 3	608	0.9341	894
Without Rule 4	526	0.9082	2307

(c) Null-Model and Alternative-Topology Comparisons

Configuration	Modularity
Observed topology	0.9045
Random graph	0.3873 ± 0.0022
Configuration model	0.3091 ± 0.0020
Exact degree-preserving randomization	0.3015 ± 0.0022
Density-matched provider-label null topology	0.8881 ± 0.0064
Text-similarity topology	0.8753

Table 3a shows that, although the modularity remains nearly unchanged after excluding Microsoft (0.9045 vs. 0.8956), the number of cross-provider bridge relationships decreases substantially (2307 to 248). This indicates that Microsoft's large capability catalog contributes disproportionately to cross-provider connectivity. Nevertheless, the persistence of high modularity after excluding the largest provider (Q = 0.8956), together with the provider-label permutation that preserves the original provider-size distribution, indicates that the observed modular organization cannot be attributed solely to the differing scales of the provider capability catalogs. Table 3b shows that different semantic edge-construction rules play complementary roles. Removing Rule 1 substantially reduces graph connectivity and community cohesion, whereas removing Rule 3 eliminates all cross-provider bridge relationships. This outcome is expected because Rule 3 explicitly defines cross-provider semantic connectivity. Consequently, the bridge analysis characterizes

the structural consequences of the adopted semantic mapping rules rather than inferring authorization equivalence automatically. In contrast, Rule 4 has only a limited influence on the overall topology. A stricter formulation of Rule 3, requiring exact normalized-resource agreement across providers, reduced the number of cross-provider bridges from 2307 to 894, while modularity remained high ($Q = 0.9341$). This indicates that the absolute number of candidate cross-provider bridges is sensitive to the adopted matching criterion, whereas the broader community-level organization remains comparatively stable under the stricter rule. The null-model comparisons provide further evidence that the observed modular organization is not explained by graph size, density, or node-degree distribution alone. The observed modularity ($Q = 0.9045$) was substantially higher than that obtained from random graphs (0.3873 ± 0.0022), configuration-model networks (0.3091 ± 0.0020), and exact degree-preserving randomizations (0.3015 ± 0.0022). The density-matched provider-label permutation retained comparatively high modularity (0.8881 ± 0.0064), but the observed topology exceeded the permutation distribution, with only one of 200 permutations reaching an equal or higher modularity (empirical $p \approx 0.010$). The observed clustering coefficient was also higher than the permutation mean (0.5437 vs. 0.3991), while the density-matched text-similarity topology produced $Q = 0.8753$ and only 299 cross-provider edges, compared with 2307 in the observed topology. These results indicate that the reported organization cannot be attributed solely to network size, degree distribution, or provider-scale imbalance, although part of the modular structure remains influenced by the declared semantic graph-construction process.

5.2 Capability Communities

The second research question examines the capability communities identified in the topology. Community detection produced 525 communities with an observed modularity of $Q = 0.9045$. Because the topology is constructed using deterministic semantic rules, this value should not be interpreted as independent evidence of naturally emerging authorization categories. However, the substantially lower modularity obtained from the random, configuration-model, and exact degree-preserving null networks indicates that graph size, density, and degree distribution alone do not explain the observed modular organization. The comparatively high modularity retained under density-matched provider-label permutation further indicates that part of the community structure remains influenced by the semantic graph-construction process. Accordingly, the detected communities are interpreted as structural properties of the declared semantic representation rather than independently discovered latent authorization categories. Table 4 summarizes the largest detected communities.

Table 4: Largest semantic capability communities.

Community ID	Scopes	Dominant Category	Providers
C320	117	Organization	Multiple
C11	77	Identity	Multiple
C26	68	Communication	Multiple
C240	53	Content	Multiple
C3	52	Organization	Multiple

The largest communities are consistently centered on organization management, identity services, communication functions, and content management, indicating that these authorization domains form the principal structural building blocks of the constructed topology. Several communities include capabilities originating from multiple providers, suggesting that different OAuth ecosystems independently

expose semantically similar authorization concepts despite using provider-specific naming conventions. This observation supports the usefulness of the proposed normalization framework for revealing cross-provider structural commonality while preserving provider-specific capability definitions.

5.3 Central Capabilities

To characterize the concentration aspect of RQ4, this subsection examines structurally important capabilities. PageRank and degree analysis identify a relatively small set of capabilities occupying central positions within the topology. These capabilities are mainly related to organization resources, identity management, and content access. Fig. 3 illustrates the highest ranked capabilities, while Table 5 summarizes the five most central capabilities identified by the analysis.

Table 5: Top central OAuth capabilities.

Capability	Provider	Degree	PageRank
team:read	Slack	99	0.0117
admin:org	GitHub	30	0.0044
admin:org_hook	GitHub	30	0.0044
user.addresses.read	Google	61	0.0022
user.gender.read	Google	61	0.0022

The highest ranked capabilities are associated with organization management, identity services, and content access. These capabilities occupy central positions within the semantic topology and connect multiple capability groups. Together with the 514 connected components reported in the topology analysis, these findings characterize the concentration and fragmentation examined in RQ4. The reported centrality values describe structural positions within the topology and should not be interpreted as operational privilege or deployment frequency.

5.4 Cross-Provider Capability Bridges

Table 6 summarizes the dominant cross-provider bridge categories and the corresponding research question findings.

The constructed topology contains 2307 cross-provider bridge relationships, primarily spanning organization, identity, communication, and content categories. These bridges represent candidate semantic correspondences under the adopted mapping rules rather than authorization equivalence. For example, Google admin.datatransfer is structurally linked to Microsoft's AdministrativeUnit.Read.All and AdministrativeUnit.ReadWrite. All through shared administration-related semantics, while Slack file capabilities form bridges with semantically related Dropbox content-management capabilities. The Rule 3 sensitivity analysis confirms that the absolute bridge count is construction-dependent, decreasing from 2307 to 894 under the stricter exact-resource criterion. Accordingly, the contribution lies in identifying where cross-provider semantic relationships occur and which authorization domains they connect, rather than treating the bridge count as an invariant property of OAuth ecosystems.

Table 6: Cross-provider bridge categories and research question findings.

(a) Bridge Categories	
Category	**Structural Observation**
Total bridge relationships	2307
Organization	Highest concentration of cross-provider semantic bridges
Identity	Common identity-management semantics across providers
Communication	Shared collaboration and messaging capabilities
Content	Semantically related file and content management capabilities
(b) Research Question Findings	
RQ	**Main Finding**
RQ1	OAuth capabilities form sparse semantic regions
RQ2	525 semantic communities were identified
RQ3	2307 rule-dependent cross-provider semantic bridges were identified
RQ4	Centrality is concentrated in organization and identity capabilities within a topology of 514 components

5.5 Practical Implications

The proposed topology provides a unified structural view of OAuth capabilities across multiple providers, supporting capability inventory management, authorization auditing, semantic scope comparison, and cross-provider analysis. It can assist security analysts during least-privilege reviews and cloud migration by identifying semantically related authorization scopes requiring manual review. Because the topology is generated from publicly documented capabilities, it can be updated and reproduced as authorization platforms evolve.

5.6 Comparison with Previous Studies

Previous studies primarily examined OAuth security, privacy, implementation, authorization management, and semantic scope organization [3,5,7–9,11,19,22]. In contrast, this study analyzes the structural organization of a unified cross-provider semantic capability network. Table 7 summarizes this distinction.

Table 7: Comparison with representative OAuth studies.

Study Focus	Security	Privacy	Scope Management	Cross-Provider Topology
OAuth security studies	✓	–	–	–
OAuth privacy studies	–	✓	–	–
OAuth deployment studies	✓	–	✓	–
Semantic scope management studies	–	–	✓	–
This work	–	–	✓	✓

The proposed analysis complements previous OAuth research by integrating semantic normalization, cross-provider capability representation, and graph-based structural analysis within a single reproducible framework. Unlike prior studies that primarily examined OAuth security, deployment behavior, privacy,

or provider-specific scope management, this work focuses on the structural organization of authorization capabilities across heterogeneous providers. Consequently, the contribution is methodological rather than protocol-centric, providing a reproducible structural representation that can support future authorization analysis and cross-provider capability comparison.

5.7 Future Research Directions

The proposed semantic topology provides a foundation for future work incorporating additional authorization providers, temporal capability evolution, and operational authorization data. These extensions may evaluate how the reported structural patterns generalize beyond static, documentation-based capability definitions.

6 Limitations

This study has several limitations. First, the capability dictionary is constructed from publicly available OAuth documentation provided by major authorization platforms. Consequently, the topology represents documented capabilities and may exclude provider-specific internal permissions or deprecated scopes. The documentation-based normalization consistency audit covered a stratified sample of 30 capabilities and did not include an independent second annotator or inter-annotator agreement analysis. Consequently, normalization reliability beyond the audited sample remains a limitation of the present study. Second, semantic relationships are generated using deterministic normalization and mapping rules. The resulting edges describe structural similarity between capabilities rather than observed authorization requests or application behavior. Third, the topology is independent of deployment environments and does not consider application popularity, user activity, authorization frequency, or runtime access patterns; therefore, the reported structural measures should not be interpreted as operational risk indicators. Fourth, the community structure and reported structural observations depend partly on the adopted semantic edge-construction rules. The robustness analyses show that this dependence varies across structural measures: high modular organization persisted under the stricter Rule 3 formulation, whereas the absolute number of cross-provider bridges decreased from 2307 to 894. Accordingly, bridge counts should be interpreted as rule-dependent structural measurements rather than invariant properties of OAuth ecosystems. Finally, this work focuses on structural organization and cross-provider capability relationships. It does not evaluate protocol vulnerabilities, privilege escalation attacks, or authorization policy enforcement. Despite these limitations, the proposed methodology provides a reproducible structural representation of OAuth authorization ecosystems and establishes a foundation for future studies combining semantic topology with operational authorization data. Future work should further evaluate the proposed framework using additional authorization providers, alternative semantic normalization strategies, and operational authorization datasets to assess the generalizability of the reported structural observations.

7 Conclusion

This study presented a structural analysis of OAuth authorization capabilities collected from multiple providers. A unified capability dictionary was constructed and transformed into a semantic topology representing cross-provider authorization relationships. Network analysis identified multiple capability communities together with central capabilities and structural bridges connecting different authorization domains. The resulting topology shows that OAuth capabilities form organized semantic structures rather than uniformly connected authorization spaces. The proposed methodology relies on publicly available capability definitions and reproducible semantic mapping rules. As a result, the complete analysis can be repeated and extended as authorization ecosystems evolve. The expanded robustness analyses show that

the observed modular organization cannot be explained by graph size, density, or node-degree distribution alone. At the same time, provider-label permutation, text-similarity comparison, and alternative Rule 3 analysis show that some structural quantities remain dependent on semantic construction choices, particularly the absolute number of cross-provider bridges. Future work may extend this topology by incorporating additional authorization providers and dynamic capability relationships.

Acknowledgement: Not applicable.

Funding Statement: The author received no specific funding for this study.

Availability of Data and Materials: All datasets, source code, network analysis outputs, and supplementary materials are publicly available at: https://github.com/maryamalmarwani2015-hub/oauth-scope-semantic-topology/tree/main. The repository contains the unified capability dictionary, semantic topology, analysis scripts, generated experimental results, and the complete Supplementary Material describing the normalization rules, category inventories, representative and ambiguous mapping cases, topology-construction rules, and released-artifact audit tables.

Ethics Approval: Not applicable.

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

Supplementary Materials: The supplementary material is available online at https://www.techscience.com/doi/10.32604/cmc.2026.086887/s1.

References

1. Luo K, Wang X, Fung PHA, Lau WC, Lecomte J. Universal cross-app attacks: exploiting and securing OAuth 2.0 in integration platforms. In: 34th USENIX Security Symposium (USENIX Security 25). Seattle, WA, USA: USENIX Association; 2025. p. 3221–38.
2. Chen EY, Pei Y, Chen S, Tian Y, Kotcher R, Tague P. OAuth demystified for mobile application developers. In: Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security; 2014 Nov 3–7; Scottsdale, AZ, USA. New York, NY, USA: ACM; 2014. p. 892–903. doi:10.1145/2660267.2660323.
3. Sun ST, Beznosov K. The devil is in the (implementation) details: an empirical analysis of OAuth SSO systems. In: Proceedings of the 2012 ACM Conference on Computer and Communications Security (CCS '12); 2012 Oct 16–18; Raleigh, NC, USA. New York, NY, USA: ACM; 2012. p. 378–90.
4. Sadqi Y, Belfaik Y, Safi S. Web OAuth-based SSO systems security. In: Proceedings of the 3rd International Conference on Networking, Information Systems & Security; 2020 Mar 30–Apr 2; Marrakech, Morocco. p. 1–7.
5. Suzic B, Prünster B, Ziegler D. On the structure and authorization management of RESTful web services. In: Proceedings of the 33rd Annual ACM Symposium on Applied Computing; 2018 Apr 9–13; Pau, France. New York, NY, USA: ACM; 2018. p. 1716–24. doi:10.1145/3167132.3167315.
6. Philippaerts P, Preuveneers D, Joosen W. OAuch: exploring security compliance in the OAuth 2.0 ecosystem. In: Proceedings of the 25th International Symposium on Research in Attacks, Intrusions and Defenses; 2022 Oct 26–28; Limassol, Cyprus. New York, NY, USA: ACM; 2022. p. 460–81. doi:10.1145/3545948.3545955.
7. Dimova Y, Van Goethem T, Joosen W. Everybody's looking for SSOmething: a large-scale evaluation on the privacy of OAuth authentication on the web. Proc Priv Enhancing Technol. 2023;2023(4):452–67. doi:10.56553/popets-2023-0119.
8. Zuo C, Zhao Q, Lin Z. AUTHSCOPE: towards automatic discovery of vulnerable authorizations in online services. In: Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security; 2017 Oct 30–Nov 3; Dallas, TX, USA. New York, NY, USA: Association for Computing Machinery; 2017. p. 799–813. doi:10.1145/3133956.3134089.
9. Morkonda SG, Chiasson S, Oorschot PC. Influences of displaying permission-related information on web single sign-on login decisions. Comput Secur. 2024;139(5):103666. doi:10.1016/j.cose.2023.103666.

10. Singh J, Chaudhary NK. OAuth 2.0: architectural design augmentation for mitigation of common security vulnerabilities. J Inf Secur Appl. 2022;65(1):103091. doi:10.1016/j.jisa.2021.103091.
11. Suzic B, Reiter A, Marsalek A. Structuring the scope: enabling adaptive and multilateral authorization management. In: Proceedings of the 2017 IEEE Conference on Communications and Network Security (CNS); 2017 Oct 9–11; Las Vegas, NV, USA. New York, NY, USA: IEEE; 2017. p. 522–30. doi:10.1109/CNS.2017.8228700.
12. Lancichinetti A, Kivelä M, Saramäki J, Fortunato S. Characterizing the community structure of complex networks. PLoS One. 2010;5(8):e11976. doi:10.1371/journal.pone.0011976.
13. Yang J, Leskovec J. Structure and overlaps of ground-truth communities in networks. ACM Trans Intell Syst Technol. 2014;5(2):1–35. doi:10.1145/2594454.
14. Kikas R, Gousios G, Dumas M, Pfahl D. Structure and evolution of package dependency networks. In: Proceedings of the 2017 IEEE/ACM 14th International Conference on Mining Software Repositories (MSR); 2017 May 20–21; Buenos Aires, Argentina. New York, NY, USA: IEEE; 2017. p. 102–12. doi:10.1109/MSR.2017.55.
15. Fortunato S, Hric D. Community detection in networks: a user guide. Phys Rep. 2016;659(12):1–44. doi:10.1016/j.physrep.2016.09.002.
16. Bansal C, Bhargavan K, Delignat-Lavaud A, Maffeis S. Discovering concrete attacks on website authorization by formal analysis. J Comput Secur. 2014;22(4):601–57. doi:10.1109/csf.2012.27.
17. Shernan E, Carter H, Tian D, Traynor P, Butler K. More guidelines than rules: CSRF vulnerabilities from noncompliant OAuth 2.0 implementations. Cham, Switzerland: Springer International Publishing; 2015. p. 239–60. doi:10.1007/978-3-319-20550-2_13.
18. Yang R, Li G, Lau W, Zhang K, Hu P. Model-based security testing: an empirical study on OAuth 2.0 implementations. In: Proceedings of the 11th ACM Asia Conference on Computer and Communications Security (AsiaCCS '16); 2016 May 30–Jun 3; Xi'an, China. New York, NY, USA: ACM; 2016. p. 651–62.
19. Lodderstedt T, Bradley J, Labunets A, Fett D. Best current practice for OAuth 2.0 security. Wilmington, DE, USA: RFC Editor; 2025. RFC 9700. doi:10.17487/rfc9700.
20. Wang R, Zhou Y, Chen S, Qadeer S, Evans D, Gurevich Y. Explicating SDKs: uncovering assumptions underlying secure authentication and authorization. In: Proceedings of the 22nd USENIX Security Symposium (USENIX Security 13); 2013 Aug 14–16; Washington, DC, USA. Seattle, WA, USA: USENIX Association; 2013. p. 399–414.
21. Glöckler J, Sedlmeir J, Frank M, Fridgen G. A systematic review of identity and access management requirements in enterprises and potential contributions of self-sovereign identity. Bus Inf Syst Eng. 2024;66(4):421–40. doi:10.1007/s12599-023-00830-x.
22. Cao L, Meng L, Stefan D, Fernandes E. Stateful least privilege authorization for the cloud. In: Proceedings of the 33rd USENIX Conference on Security Symposium, SEC '24; 2024 Aug 14–16; Philadelphia, PA, USA. Seattle, WA, USA: USENIX Association; 2024.
23. Leskovec J, Lang KJ, Dasgupta A, Mahoney MW. Statistical properties of community structure in large social and information networks. In: Proceedings of the 17th International Conference on World Wide Web; 2008 Apr 21–25; Beijing, China. New York, NY, USA: ACM; 2008. p. 695–704. doi:10.1145/1367497.1367591.
24. Hernández R, Gutiérrez I, Castro J. Social network analysis: a novel paradigm for improving community detection. Int J Comput Intell Syst. 2025;18(1):87. doi:10.1007/s44196-025-00812-9.
25. Blondel VD, Guillaume JL, Lambiotte R, Lefebvre E. Fast unfolding of communities in large networks. J Stat Mech Theory Exp. 2008;2008(10):P10008. doi:10.1088/1742-5468/2008/10/P10008.