ADVANCES IN KNOWLEDGE ORGANIZATION
Proceedings of International ISKO Conference - ISSN: 0938-5495 - 2026 - São Paulo

Between Censorship and Curation: The Legitimacy Dilemma of Knowledge Organization in Online Health Forums

Mouheb Mehdoui

Gériico Laboratory, University of Lille, France

Amel Fraisse

Gériico Laboratory, University of Lille, France

Widad Mustafa El Hadi

Gériico Laboratory, University of Lille, France

Mounir Zrigui

RLANTIS Laboratory, University of Monastir, Tunisia

Abstract

Online health communities (OHCs) are vital for peer support, yet their governance is undertheorized. Rules are not neutral but enact epistemic governance, privileging certain voices. Analysis of 738 rules across 117 Reddit health communities reveals a core asymmetry: high health literacy (HHL) contributions are legitimized, while low health literacy (LHL) narratives are contained—termed Governance-Induced Epistemic Silencing. This follows an Epistemic Risk-Calibration Model: chronic/contested illnesses adopt restrictive models; infectious/general ones adopt integrative approaches, creating a Participatory-Curatorial Divide and Diagnostic Determinism. The framework explains how platforms institutionalize authority and marginalize patients. Bridging infrastructures are advocated over prohibitive governance, and rules are viewed as dynamic instruments for more just digital health ecosystems.

Keywords: Knowledge Organization (KO); Online Health Communities (OHCs); Epistemic Governance; Health Literacy; Platform Moderation; Epistemic Justice; Censorship and Curation; Reddit; Socio-technical Systems; Legitimacy Dilemma; Infrastructure

1.0 Introduction

The rapid, crisis-driven shift to digital public health engagement has offered powerful tools for information dissemination and disease control, but it also introduces systemic risks to privacy, autonomy, and fairness. A primary locus of this engagement is the online health community (OHC), where individuals navigate the promises and perils of digital health. As demonstrated in Nölke et al. (2015), OHCs are vital sources of peer support and information, yet their internal governance remains underexamined. Each community establishes formal rules — its" Rules,"" Guidelines," or "Community Standards"

— that function as instruments of Knowledge Organization (KO). These rules determine what information is legitimate, how it is categorized, and who may contribute, thereby creating a pre-emptive governance structure that shapes all subsequent interaction. Moderators establish rules to prevent misinformation, hate speech, and bullying, and they may add flags, delete certain terms, or restrict users to preserve a professional environment. However, the boundary between organizing knowledge and silencing voices is often thin; the same rules that legitimize contributions can also pre-emptively exclude them.

This creates what we call a legitimacy dilemma: a core governance conflict between validating user contributions and managing informational risk. The resolution of this dilemma varies systematically by disease context, placing communities on an equity spectrum from facilitative curation to restrictive censorship. In practice, stringent KO rules can transition from organizing knowledge to enacting censorship, thereby creating unequal legitimacy in accessing medical information and exacerbating existing health disparities. Despite the centrality of these rule systems to millions of patients and caregivers, the governance of knowledge in OHCs remains undertheorized as a phenomenon of Knowledge Organization. Research on platform moderation has focused overwhelmingly on commercial content policies and misinformation at the platform level, obscuring the community-specific, pre-emptive KO infrastructures that quietly constitute epistemic participation (Gillespie 2018). At the same time, KO scholarship has largely remained anchored in concept analysis and bibliographic contexts, leaving unaddressed how classification, categorization, and warrant operate in peer-produced health spaces where the boundaries between “document” and “person,” “subject heading” and “user flair,” and “catalog” and “megathread” have collapsed (Dahlberg 1978; Hjørland 2008).

This gap has tangible consequences. Patients — especially those with conditions requiring substantial support — may feel abandoned, restricted, or silenced on platforms where they expect to interact freely. When community KO rules systematically privilege biomedical citation over lived experience, they risk reproducing the hierarchies of credibility that drive patients to peer forums. When infrastructural tools such as flair and structured templates are deployed without reflexive attention to whose knowledge they amplify, they can transform supportive spaces into sites of testimonial silencing (Fricker 2007). And when the design of these systems varies systematically by disease type — imposing stricter epistemic gatekeeping on contested, chronic, or stigmatized conditions — the result is a form of diagnostic discrimination encoded in metadata and rule text. We therefore ask four research questions. First (RQ1), how do OHCs regulate users and contributions through KO rules? Second (RQ2), what infrastructural governance mechanisms do they employ? Third (RQ3), how does disease moderate these governance choices? Fourth (RQ4), what emergent governance archetypes can be identified?

The remainder of this paper is structured as follows. Section 2 establishes the theoretical foundation, reviewing digital health literacy and epistemic (in)justice. Section 3 presents our analytical framework. Section 4 details our methodology. Section 5 reports empirical findings. Section 6 discusses implications through a novel Epistemic Risk-Calibration

Framework. Section 7 concludes by calling for conscious redesigning socio-technical systems to foster equitable epistemic participation.

2.0 Literature Review and Theoretical Framework

Knowledge is never neutral but is always shaped by the actors who interact with, produce, and define the directions of a domain, and remains deeply influenced by intersecting theories of organization, literacy, and justice (Berger and Luckmann 1966; Foucault 1980; Haraway 1988; Fricker 2007). To explore this intersection, we draw on three core literatures that together demonstrate how Knowledge Organization rules function as sociotechnical tools of epistemic control (Olson 2002). At its core, KO is the discipline that studies the nature, structure, and representation of concepts and their relations—how knowledge domains are organized, how conceptual systems are built, and how different epistemological positions shape those systems (Dahlberg 1978; Hjørland 2008). Foundational KO scholarship centres on concept analysis, warrant (what justifies a class or category), and the integration of multiple epistemic perspectives, addressing questions of epistemology and domain analysis rather than merely the efficient retrieval of documents (Szostak et al. 2016). This conceptual understanding is essential for analyzing online health communities, where KO is enacted through community-specific mechanisms: content flairs, mandatory tags, source-validation rules, megathreads, and structured post templates. These are not cataloguing instruments but conceptual structures that implicitly define what counts as legitimate knowledge, who is authorised to create or apply categories, and how different forms of knowledge are related or segregated.

A sociotechnical perspective rejects the dichotomy between the technical architecture of a system and the social order it produces, recognizing them as mutually constitutive (Dourish 2001). In online communities, the technical infrastructure of KO—flairs, wikis, automated filters—is the means through which social and epistemic governance is enacted. These systems govern by performing three critical functions. First, they structure participation by channeling user behavior into sanctioned formats and spaces. Second, they establish authority by creating explicit or implicit hierarchies of credibility, such as rules that prioritize peer-reviewed citations over personal anecdote, thereby determining whose knowledge counts (Gillespie 2018). Third, and most fundamentally, they create legitimacy by defining the normative boundaries of the community, adjudicating what constitutes a helpful contribution versus dangerous misinformation. Consequently, analyzing a community's formal rules is synonymous with analyzing its KO system, and analyzing its KO system is an analysis of its core governance logic.

The concept of health literacy has become a crucial aspect of public health, encompassing the personal, cognitive, and social skills which determine the ability of individuals to access, understand, and use information in ways that promote and maintain good health (Nutbeam 2000).The transition of health-seeking and peer support activities to online platforms has broadened this concept into digital health literacy, which involves the skills to seek, find, understand, evaluate, and use health information from electronic sources (Norman and Skinner 2006). Within the governed spaces of online health communities,

digital health literacy translates into distinct epistemic and navigational competencies required for meaningful participation. Operationally, this encompasses the ability to navigate the platform's interface and its knowledge-organizing features. Critically, it demands the informational and critical literacy necessary to evaluate the credibility of sources and claims amidst a high volume of peer-generated content. Productively, it necessitates communicative literacy to articulate experiences and provide support in ways that align with community norms to be recognized as legitimate contributions. The present study draws upon this multidimensional understanding not to assess individual deficits, but to analyze how community governance systems anticipate and manage a spectrum of user competencies. This approach inverts a common critique of health literacy that imposes a deficit model on individuals (Zarcadoolas et al. 2005). Instead, our analysis investigates how formal governance structures are designed in response to this literacy continuum, shifting the central inquiry from user capability to systemic design: how do rules facilitate or constrain participation based on a user's projected literacy, and what are the equity implications of designing governance around literacy-based assumptions?

To evaluate the governance designs and their differential impact, we adopted Miranda Fricker's (2007) framework of epistemic (in)justice. Epistemic injustice denotes a mistake done to someone specifically in their capacity as a knower, manifesting in two primary forms. Testimonial injustice occurs when a speaker is afforded less credibility than they deserve due to identity prejudice, such as when a patient's narrative is systematically discounted because they lack formal medical authority. Hermeneutical injustice arises from a structural gap in shared interpretative resources, which leaves individuals unable to comprehend or articulate key aspects of their own social experience. In digital health communities, these epistemic struggles are recapitulated and intensified. These spaces are contested arenas where the traditional authority of biomedical expertise interacts with, and is sometimes challenged by, the lived experience of patients. Consequently, the formal governance rules of a community—its KO infrastructure—function as the primary mechanism for either perpetuating or remedying these injustices. This study operationalizes Fricker's framework to critically assess community rules by asking two guiding questions: do the KO rules enact testimonial injustice by systematically amplifying or silencing certain speakers based on their presumed literacy, and does the KO infrastructure provide the hermeneutical resources that empower all users to interpret and name their reality, or does it restrict legitimate discourse to a narrow biomedical lexicon thereby enacting hermeneutical injustice? This lens provides the crucial normative dimension to the KO-Equity Spectrum introduced in our conceptual framework. Epistemic justice aligns with the participatory curation end of the spectrum, where KO tools are designed to scaffold diverse contributions and validate multiple forms of knowledge. Epistemic injustice, conversely, is the inherent risk of the pre-emptive censorship pole, where restrictive rules foreclose certain forms of speech and credibility. The legitimacy dilemma at the heart of this study is fundamentally a dilemma of epistemic justice: how can communities validate the urgent, experiential knowledge of the help-seeker without enabling harm, while also leveraging the systematized knowledge of the support-giver without instituting an exclusionary hierarchy?

3.0 Conceptual Framework

To respond to the research questions and validate the proposed hypotheses, we establish an analytical framework centered on the tension between different levels of health literacy and the expected behavior of each level, as intensified by Knowledge Organization practices. To distinguish these levels, we employ a user typology that illuminates the legitimacy dilemma — the tension between validating personal narratives and curated expertise — and enables us to map communities along a spectrum from censorship to curation. Figure 1 presents the theoretical framework integrating KO, health literacy, and epistemic (in)justice.

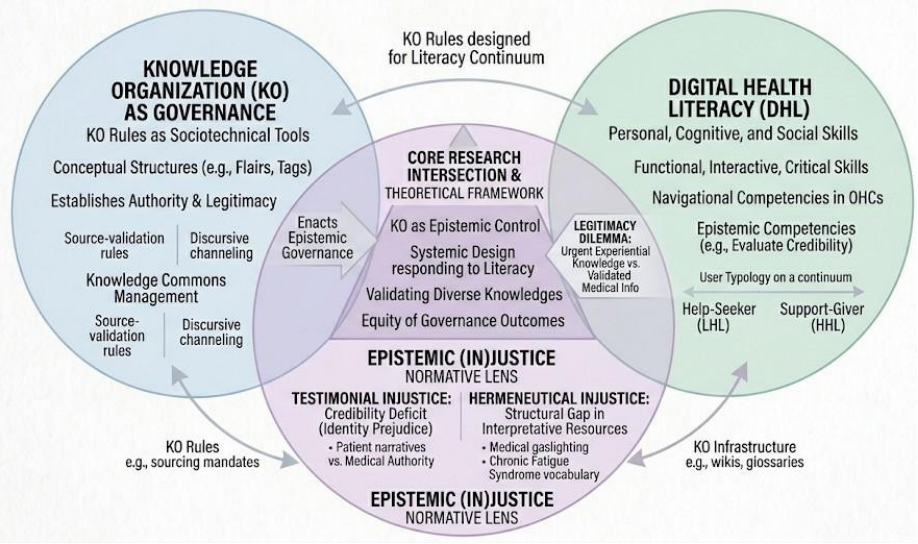


Figure 1. Theoretical framework: KO, health literacy, and epistemic (in)justice

Building directly on the theoretical manifestation of health literacy in digital contexts, the present work makes a fundamental distinction between Low-Health-Literacy and High-Health-Literacy users as analytical ideal types. These are not conceived as fixed demographic categories but as constructs designed to reveal how community governance systems anticipate, imagine, and differentially regulate participants and their contributions. The Low-Health-Literacy user is analytically characterized by a primary objective of help-seeking, with epistemic competency modeled as residing strongly in personal, experiential narrative. Consequently, their archetypal contribution style includes personal stories, direct questions, and requests for shared experience. The High-Health-Literacy user is analytically characterized by a primary objective of support-giving or expertise-sharing, with competency modeled as encompassing greater fluency in biomedical discourse, information validation, and sophisticated use of the community's knowledge-organizing infrastructure. Their archetypal contribution style involves giving advice, sharing curated information, and validating or moderating the claims of others. These ideal-typical constructs link assumed competencies to anticipated platform behaviors. We adopted this binary typology as a deliberate analytical simplification; while models of health literacy acknowledge a spectrum of competencies, our framework

focuses on the archetypal contribution styles that community rules are designed to regulate. A user with moderate literacy is not statically categorized but is anticipated to oscillate between these behavioral modes based on context. Consequently, the platform's governance logic is triggered by the mode of contribution: when a user posts a personal narrative, Low-Health-Literacy-oriented rules apply; when they share a vetted citation, High-Health-Literacy-oriented rules engage. The typology therefore models the behavioral modes that KO rules channel, validate, or restrict, enabling a parsimonious analysis of systemic responses to the epistemic tension between narrative and expertise.

The interaction between defined archetypal users and a community's knowledge organization system creates a constitutive tension termed the legitimacy dilemma. This fundamental governance challenge arises because the platform's rules must simultaneously validate two distinct and potentially contradictory forms of participation. On one hand, the system must legitimize the essential help-seeking of Low-Health-Literacy users, whose personal narratives provide the foundational layer of peer support and lived experience that attracts and sustains community engagement. However, this validation inherently accommodates the risk of unintentional misinformation, as personal anecdotes may be non-generalizable, emotionally compelling but inaccurate, or mistakenly presented as medical fact. On the other hand, the system must also legitimize the valuable support-giving of High-Health-Literacy users, whose advice and information curation elevate the community's utility and credibility. This validation inherently accommodates the risk of unauthorized or harmful advice, where even well-intentioned information may cross into prescriptive guidance, oversimplify complex issues, or originate from sources without appropriate authority. A community's Knowledge Organization rules are therefore a strategic and inherently political resolution to this inescapable trade-off. Crucially, this resolution is never merely technical; it carries significant normative weight and consequences for equity. A common enforcement action — banning content deemed to contain misinformation or an incorrect diagnosis — can function as a form of censorship that disproportionately impacts Low-Health-Literacy users. Since their contributions are rooted in personal narrative rather than formal medical knowledge, they may unknowingly share inaccurate information. Penalizing this unknowing error effectively sanctions individuals for their lack of expertise, potentially silencing vulnerable voices seeking help. Conversely, automatically allowing content from High-Health-Literacy users because it appears factually correct confers automatic legitimacy upon them, institutionalizing their epistemic authority within the community. This dynamic creates a systemic inequity: the rules themselves can automatically privilege the contributions of the health-literate while marginalizing those of the less literate, thereby reinforcing the very knowledge gaps the community might aim to bridge. The primary analytical task is to examine how these rules resolve the dilemma — whether through a restrictive approach that broadly limits certain contribution types to minimize risk, a channeling approach that carefully segregates and labels narratives and advice into designated spaces, or a synthetic approach that actively designs for productive interaction between experiential and informational contributions. Each resolution implicitly prioritizes one form of legitimacy while managing the other's risks, ultimately defining the community's epistemic character and its normative vision of what constitutes trustworthy knowledge and legitimate participation.

To systematically interpret and categorize the strategic resolutions to the legitimacy dilemma, this analysis employs the KO-Equity Spectrum. This heuristic maps the observable continuum of governance logics onto a single axis, defined by their operational principles and their direct implications for epistemic equity. The spectrum is anchored by two opposing poles that represent fundamentally different approaches to managing risk and participation. At the Pre-emptive Censorship Pole, governance is enacted primarily through broad prohibitions designed to eliminate risk ex-ante, such as “No personal anecdotes,” “No medical advice,” or “Do not ask for diagnoses.” This restrictive logic prioritizes safety and liability mitigation by constraining the scope of allowable speech. Analytically, this approach often disproportionately silences the primary mode of Low-Health-Literacy contribution, potentially enacting testimonial injustice by systematically devaluing experiential knowledge as inherently unreliable. At the Participatory Curation Pole, governance is enacted through organizational tools designed to scaffold and guide participation, such as mandatory diagnostic flairs, dedicated megathreads, or structured wiki templates for condition summaries. This facilitative logic uses KO not to prohibit but to shape diverse contributions into a legible and organized commons. It acknowledges risk but seeks to manage it through structure rather than silence, thereby promoting epistemic justice by creating structured, inclusive pathways for both help-seeking and expertise-sharing. A community’s dominant position on this spectrum crystallizes its fundamental governance logic and its implicit stance on the legitimacy dilemma. Mapping and comparing these positions provides the primary analytical tool for this comparative study, allowing for a nuanced evaluation that directly interrogates the critical distinction between knowledge organization as a form of legitimate, ordering governance and its potential erosion into exclusionary censorship. This spectrum reflects competing epistemological positions within KO: the pre-emptive censorship pole corresponds to a narrow scientific warrant and a positivist epistemology that privileges decontextualized, generalizable knowledge, while the participatory curation pole aligns with a pluralist epistemology and multiple warrants— including user warrant and cultural warrant — that recognize lived experience as a legitimate basis for categorization according to Hjørland (2008), Szostak et al. (2016).

4.0 Methodology

Building on the conceptual framework draft drawn previously, the core methodology adopts a systematic, comparative qualitative document analysis of the formal governance texts of online health communities, using Reddit as a principal site of investigation. A qualitative approach is essential to interpret the meaning, nuance, and underlying governance logic embedded within community rules, moving beyond quantification to understand how knowledge organization principles are articulated as instruments of social order. The methodology is explicitly designed as a document analysis because the research questions target the declared, constitutional architecture of communities; their publicly posted rules, wikis, and guidelines constitute the primary data source, functioning as the” institutional arrangements” fiesler of the knowledge commons. The design is comparative and systematic to identify coherent patterns, tensions, and variations across a large

sample of Reddit communities (N=109 subreddits), enabling a robust mapping of communities onto the KO-Equity Spectrum and a rigorous test of the analytical framework's explanatory power across diverse health contexts. This approach directly operationalizes the theoretical framing developed previously, treating rules as governance texts and their embedded knowledge organization principles as the core sociotechnical system under investigation.

4.1 Data Collection

The sample for this study was drawn exclusively from Reddit[1], a major platform hosting thousands of health-related communities known as "subreddits." Each community is primarily dedicated to a specific health condition or concern (e.g., cancer, ADHD, bipolar disorder). A key feature of Reddit's architecture is that platform-wide content policies are minimal; the primary governance and rule-setting occur at the level of individual subreddits through their moderators. This decentralized structure provides a unique and controlled research environment: it effectively holds the platform-level variable constant, allowing for a focused comparative analysis of community-level governance and knowledge organization across a diverse set of health contexts. In total, we identified 1,138 health-related subreddits from a comprehensive scan of the platform. From this initial corpus, we employed a purposive stratified sampling strategy to construct a diverse and analytically relevant sample. The sampling strategy employed purposive stratification to ensure representation across key disease and health categories, while also accounting for variation in community size and activity—capturing a spectrum from large, general forums to smaller, niche support groups. All selected subreddits were required to meet strict inclusion criteria: English-language, explicit focus on peer-to-peer health discussion or support, and the presence of a publicly accessible, substantive set of formal rules. The resulting final sample comprised 117 communities, representing four broad categories: specific somatic conditions (e.g., cancer-related, diabetes, asthma), mental health (e.g., anxiety, ADHD), chronic or contested illnesses (e.g., long Covid, chronic illness), and broad health domains (e.g., health, medicine, public health). From these communities, we programmatically extracted all formal governance texts, yielding a corpus of 1,545 elaborative rule statements drawn from wikis, FAQs, and pinned posts. Through systematic analysis, these were consolidated into 738 core, discrete rule statements that represent the fundamental governance principles of each Community of Reddit (https://www.reddit.com/), a major platform.

Table 1 Reconstructed codebook of master categories and emergent codes used in the analysis of OHC governance rules

Master Category	**Definition / Focus**	**Pre-defined Coding Dimensions / Keywords**	**Emergent Sub-codes**	**Illustrative Rule Indicators**
Legitimacy	Rules that define what	legitimacy, authority,	Professional	"No medical advice"; "No

Master Category	Definition / Focus	Pre-defined Coding Dimensions / Keywords	Emergent Sub-codes	Illustrative Rule Indicators
	counts as legitimate knowledge, credibility, authority, and acceptable contribution; manage misinformation and risk.	credibility, evidence, misinformation, medical advice, diagnosis	boundary	diagnosing"; "Cite peer-reviewed sources"; credentialing of HHL contributions; conduct regulation
Literacy	Rules that anticipate and regulate different health-literacy competencies and contribution styles (LHL vs. HHL).	health literacy, LHL, HHL, narrative, experience, advice, comprehension	Trauma sensitivity	Content warnings; disclaimers; structured templates for personal stories; plain-language expectations; channeling narrative vs. expertise
Infrastructure	Socio-technical KO tools that structure participation and organize community knowledge.	flair, tags, templates, megathreads, wiki, FAQ, automation, knowledge base	Knowledge base integration	Structured formats; flair systems; megathreads; community wikis; automated resource linking (often absent)
Equity	Normative dimension assessing inclusion, access, and epistemic (in)justice; positions on the	equity, justice, inclusion, access, censorship, curation, silencing	Crisis limitation	Prohibitive vs. facilitative rules; crisis/emergency restrictions; rules that silence LHL narratives or support HHL

Master Category	Definition / Focus	Pre-defined Coding Dimensions / Keywords	Emergent Sub-codes	Illustrative Rule Indicators
	censorship–curation spectrum.			curation

4.2 Data Analysis and Ethical Considerations

Data analysis followed an iterative, hybrid deductive-inductive qualitative coding process. The unit of analysis was the discrete rule statement (n=738), extracted from both primary rules and their sub-rules. Analysis was conducted using a custom qualitative analysis framework specifically designed for this study to manage, code, and analyze the textual corpus systematically. The analytical process began deductively with a provisional codebook structured around the four master categories derived from the conceptual framework: Legitimacy, Literacy, Infrastructure, and Equity, as shown in Table 1. Initial coding applied pre-defined keyword patterns corresponding to each category, establishing a foundational coding layer aligned with the theoretical constructs. Subsequently, three cycles of inductive, data-driven coding were performed, involving line-by-line examination of the rule statements to capture emergent themes and sub-codes not anticipated by the initial framework. Examples of emergent codes included professional boundary under Legitimacy, trauma sensitivity under Literacy, knowledge base integration under Infrastructure, and crisis limitation under Equity. The codebook was refined iteratively after each cycle, with a threshold applied where emergent codes appearing in fewer than five percent of rule statements were re-evaluated for salience. This iterative refinement between the data, codes, and conceptual framework ensured analytical rigor, consistency, and theoretical saturation.

The analysis of governance structures within online health communities necessitates consideration of the researchers' positionality and relevant ethical parameters. This study adopts the position of an external observer, analyzing publicly articulated community rules rather than participating in or evaluating internal community dynamics. While this provides necessary analytical distance, it also means the analysis is limited to formal, textual representations of governance and does not capture the nuanced, lived experiences of moderation or unwritten community norms. The conceptual framework employed, centred on the tension between censorship and curation, inevitably shapes the interpretation of the data, foregrounding certain governance aspects while backgrounding others. A reflexive approach was maintained throughout to acknowledge this interpretive lens. Ethically, this study was conducted using only publicly available community rules, which are accessible to any visitor without authentication or data collection from individual users. No private communications, user-generated content, or personally identifiable information was accessed or analyzed. The research focuses on aggregate patterns and comparative

governance models, avoiding the identification or critique of specific communities. It is important to note that the analysis of written rules reveals stated governance ideals and formal structures; it cannot account for the practical implementation of these rules or the informal cultural practices that also constitute community governance. The findings should therefore be interpreted as an analysis of normative, aspirational governance as documented in community guidelines, rather than a comprehensive account of operational community management.

5.0 Findings

Our analysis reveals a consistent and measurable asymmetry in how communities regulate contributions based on health literacy. Low-health-literacy narratives are often contained and silenced, while high-health-literacy contributions are credentialed and infrastructurally supported. The comparative analysis across 117 online health communities uncovers that asymmetry in how contributions related to low- and high-health-literacy types are regulated, thereby enacting the legitimacy dilemma. That difference appears at the most basic structural level: while most communities (70.9%, n=83) have rules explicitly designed to shape high-health-literacy-aligned contributions, fewer than half (48.7%, n=57) have rules specifically targeting low-health-literacy contributions. More significantly, governance attention varies considerably; communities use, on average, 2.77 rules for high-health-literacy users compared to 1.72 for low-health-literacy users, yielding an overall rule ratio of 0.62 low-health-literacy to high-health-literacy rules.

That volume imbalance is further highlighted by qualitative differences in governance logic. Analyzing rule stringency reveals that low-health-literacy contributions are mostly managed through prohibitive (48.5% of low-health-literacy-coded rules) and containment strategies, including common mandates like “This is not a medical community,” which systematically nullify the epistemic and prescriptive aspects of personal experience by banning requests for diagnosis or medical advice, while also directing anecdotal sharing into separate spaces or requiring content warnings and disclaimers. Conversely, high-health-literacy contributions are primarily governed through credentialing (33.1% of high-health-literacy-coded rules) and conduct regulation, establishing clear credibility hierarchies by emphasizing peer-reviewed sources and official guidelines, supporting curatorial efforts through community-maintained wikis and FAQs, and setting standards for civility while opposing overt gatekeeping. That dual framework creates a pronounced legitimacy imbalance: 46.8% of all low-health-literacy-related rules are prohibitive, while rules that actively promote high-health-literacy contributions are minimal (0.6%). The resulting Legitimacy Asymmetry Ratio of 47.00 vividly illustrates institutionalized inequality: community rules systematically impose more barriers to legitimate participation for help-seeking low-literacy users by limiting their contributions, while lowering barriers and providing infrastructure to support high-literacy users in shaping the community’s primary informational discourse. Rules not only regulate behavior but embed a normative order endorsing biomedical expertise and curatorial authority, thus institutionalizing epistemic advantages within peer support spaces. Figure 2 displays the governance asymmetry, showing that nearly 71% of communities have explicit rules shaping high-health-literacy

contributions compared to only 49% with rules targeting low-health-literacy contributions, with high-health-literacy contributions attracting more governance attention per community.

That rule-based asymmetry is practically implemented and reinforced through specific infrastructural governance tools. Communities apply normative logics with a limited technical toolkit: structured formats, flair systems, and megathreads. Structured formats — mandatory templates or required elements for posting — are the most common, found in 47.0%. The function of these tools is not neutral; each serves as a technical mechanism directly instantiating governance logics identified earlier. Structured formats act as a procedural funnel, disciplining user discourse at the point of entry by requiring specific information fields, disclaimers, or evidence citations, thereby implicitly privileging high-health-literacy-aligned contributions while formalizing and constraining the open-ended narrative expression characteristic of low-health-literacy contributions. Flair systems function as a taxonomic apparatus for classification, segregation, and filtering of content, enabling the channeling logic applied to low-literacy content by creating bounded, searchable categories for personal experiences while also empowering high-literacy users to efficiently locate and engage with content matching their expertise or curatorial interests. Megathreads represent the infrastructural pinnacle of containment, physically aggregating high-volume or sensitive low-literacy-aligned topics into a single controlled space to preserve the main forum for other discourse types.

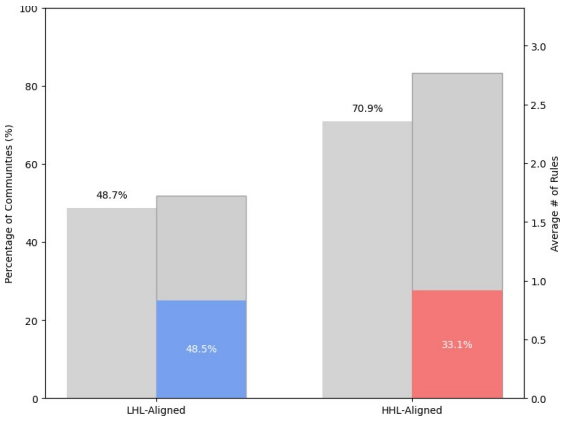


Figure 2. Asymmetry in Governance Attention and Logic for LHL vs. HHL Contributions

Crucially, the near-zero co-occurrence between these tools indicates isolated implementation, not an integrated user-centric support ecosystem. That fragmentation is compounded by the stark absence of knowledge scaffolding tools; the lack of automated resource linking or community-owned knowledge bases means the burden of information retrieval and validation falls almost entirely on real-time peer-to-peer interaction, inherently advantaging users with the literacy to efficiently find and cite credible sources. Consequently, the infrastructural landscape materializes and amplifies legitimacy

asymmetry codified in the rules, providing high-literacy users with organizational schemata to efficiently navigate and govern informational space while offering low-literacy users infrastructure primarily designed to classify, contain, and streamline moderation of their contributions. Figure 3 shows structured formats appearing in nearly half of communities, flair systems in about one-third, megathreads in only 6.8%.

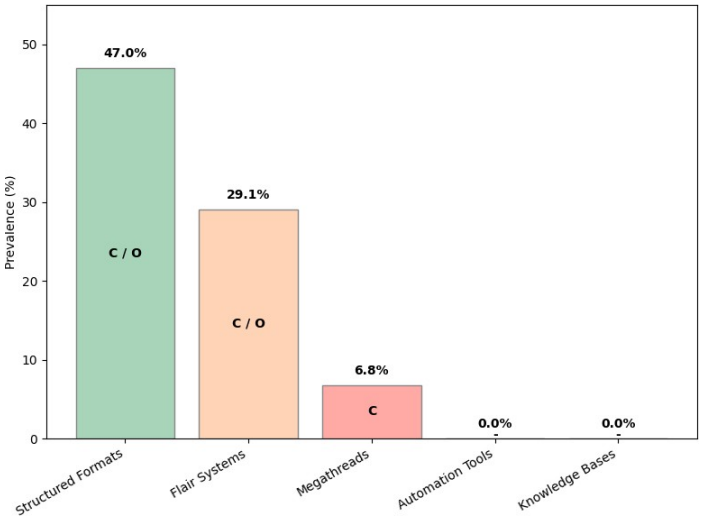


Figure 3. Prevalence and Function of Infrastructural Governance Tools

The epistemic and communicative characteristics of a community's primary health focus exert significant, systematic influence on its chosen knowledge organization model, thereby shaping the concrete instantiation of the legitimacy dilemma. Data demonstrates clear stratification of communities along the KO-EQ spectrum based on disease category. Communities organized around infectious diseases (mean spectrum position = 0.81, SD = 0.21) and general health (mean = 0.66, SD = 0.30) align strongly with the high-equity, high-curation pole, reflecting a knowledge organization paradigm designed for aggregating and disseminating relatively stable, guideline-based public health or clinical information. In stark contrast, communities focused on chronic illnesses (mean = 0.32, SD = 0.36) position decisively toward the censorship pole, indicative of a governance model prioritizing stringent informational gatekeeping and risk mitigation in contexts where medical knowledge is often evolving, contested, or burdened with high potential for harmful misinformation.

Inferential statistical analyses confirm that these governance differences are not random. A one-way ANOVA found statistically significant variation in mean spectrum positions across disease categories ($F = 4.09$, $p = 0.005$), a result robust to non-parametric evaluation (Kruskal-Wallis $H = 13.53$, $p = 0.009$). While the effect size is modest, it underscores disease context as a meaningful, though not exclusive, determinant of governance logic. A chi-square test of independence further elucidates that relationship, revealing significant association between disease category and dominant governance archetype. Examination of standardized residuals pinpoints specific organizational dispositions: chronic illness

communities significantly over-represent Censorship approaches and under-represent Curation models, pointing to a knowledge organization strategy fundamentally organized around containment — corralling epistemic risks and subjective uncertainties inherent to long-term, poorly bounded conditions through prohibitive rules and channeling mechanisms. Conversely, infectious disease communities demonstrate pronounced and significant affinity for Curation approaches and complete absence of Censorship models, with knowledge organization structured for integration and synthesis, facilitating collective vetting and architectural arrangement of authoritative information.

The centrist, variable positions of mental health and surgical communities reflect hybrid or situational models that balance these imperatives. Notably, direct comparison between aggregated high-consensus and contested illness categories did not yield significance, suggesting the salient factor is not merely formal medical consensus but the broader epistemic ecology of the condition — including diagnostic objectivity, stigma, the role of patient expertise, and typical care trajectory. Disease context functions as a powerful moderator calibrating the core trade-off between credibility and inclusivity in knowledge organization. Chronic illness forums, navigating contested evidence and high vulnerability, institutionalize knowledge organization through defensive architecture that restricts and segregates personal narrative to protect communal credibility. Infectious disease and general health forums institutionalize knowledge organization through curatorial architecture designed to collect, hierarchize, and display information. The barrier to legitimate epistemic contribution is therefore not fixed but contextually variable, systematically higher in communities where underlying medical knowledge is itself ambiguous, thereby re-inscribing broader societal legitimacy dilemmas of specific health conditions into the digital realm. Figure 4 displays infectious disease communities ranking highest toward curation (0.81), general health communities next (0.67), mental health (0.52) and surgical (0.51) communities occupying moderate middle ground, and chronic illness communities positioned closest to censorship (0.37).

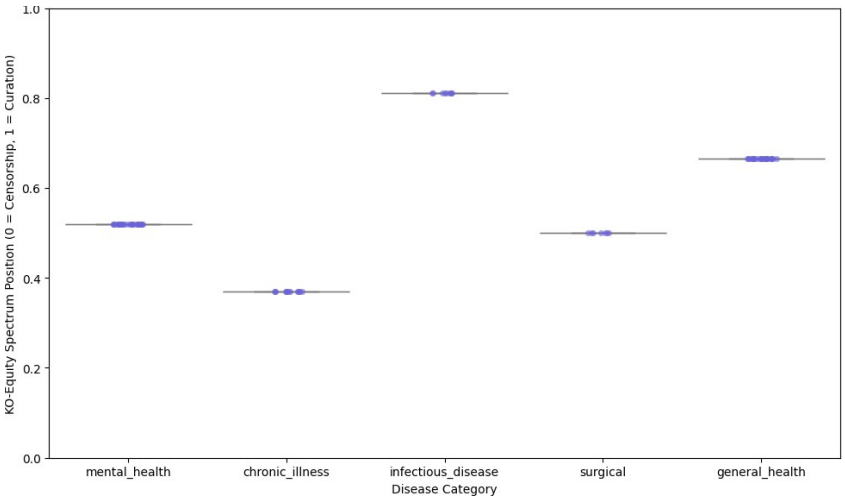


Figure 4. Stratification of Governance Models by Disease Context

To empirically map the 117 communities within the conceptual KO-EQ space and identify dominant governance archetypes, principal component and cluster analysis were conducted. The optimal clustering solution, determined by a peak silhouette score of 0.672 for six clusters, was consolidated for interpretability into four primary archetypes with a silhouette score of 0.503, still indicating reasonable and clear grouping, revealing a non-continuous landscape of governance models. Principal Component Analysis reduced three defining variables — spectrum position, censorship score, curation score — into two core components that together explain over 93%.

The four emergent clusters represent distinct strategic resolutions to the legitimacy dilemma. Strict Moderation (n=26) anchors the censorship pole, characterized by high censorship and minimal curation, functioning as a defensive, credibility-preserving model. Its antipode, Structured Community (n=39), resides at the equity pole, defined by exceptionally high curation and low censorship, prioritizing architectural integration of knowledge. Balanced Approach (n=37) and Organized Discussion (n=15) occupy the spectrum's center but differ internally: the former employs moderate levels of both censorship and curation, while the latter utilizes higher volumes of both governance actions, suggesting active management style. Statistical validation confirms robustness; one-way ANOVA reveals significant differences between clusters on all three defining variables, and post-hoc tests confirm Strict Moderation and Structured Community spectrum positions are significantly different from all others and from each other. Chi-square test of association between cluster membership and tripartite spectrum category is highly significant, with Strict Moderation overwhelmingly aligning with Censorship and Structured Community aligning with Curation. Spearman correlation matrix further validates the theoretical framework, showing strong positive correlation between spectrum position and curation score and strong negative correlation with censorship score.

That empirical mapping demonstrates the legitimacy dilemma is not resolved monolithically but through a finite set of coherent governance portfolios, with communities tending to optimize either for epistemic security through restrictive censorship or for participatory equity through active curation, with fewer achieving genuine, stable balance. Distribution across clusters provides a systematic typology for understanding how the tension between inclusivity and credibility is institutionally managed in the digital health sphere. Figure 5 visualizes these four archetypes mapped across Curation versus Censorship (horizontal axis) and Governance Intensity (vertical axis): Strict Moderation clusters near the censorship pole with high governance intensity, Structured Community at the curation end with moderate intensity, Balanced Approach occupying a wide middle band, and Organized Discussion showing high curation coupled with high governance activity.

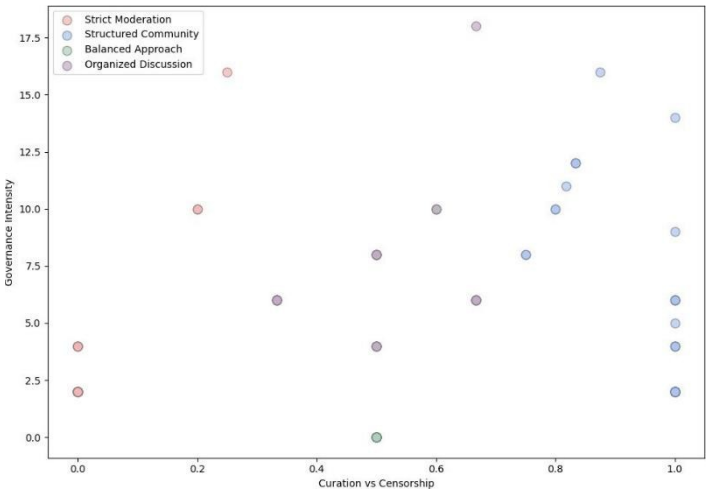


Figure 5. Governance Archetypes in the KO-Equity Space

6.0 Discussion

These findings reveal a systemic architecture of epistemic governance that calibrates rules and tools according to perceived risks across health contexts. Such governance actively constructs hierarchies of credible speech, marginalizing patient experience in spaces designed for support. The legitimacy dilemma manifests not as an abstract tension but as a coherent governance system expressed through interconnected phenomena. That system operates through multiple reinforcing mechanisms that together constitute a distinctive mode of digital knowledge regulation.

The core mechanism, Governance-Induced Epistemic Silencing, emerges from significant rule imbalance: low-health-literacy contributions face prohibitive containment, while high-health-literacy contributions receive credentialing and facilitation. That silencing is not passive but actively enacted and reinforced by a Non-Neutral Infrastructure. The limited toolkit of structured formats, flair systems, and megathreads acts as a political technology that channels personal narratives into segregated categories while simultaneously creating credentialed pathways for high-literacy users to curate and oversee discourse. Infrastructure therefore functions not as neutral scaffolding but as an active agent in shaping whose knowledge gains visibility and credibility.

That infrastructural design institutionalizes a Participatory-Curatorial Divide, assigning distinct class roles: low-literacy users contribute raw experience; high-literacy users serve as authorized curators. Such a divide entrenches health literacy disparities into fixed community positions, making literacy-based inequality structural rather than incidental. Communities effectively create two tiers of participation: one for those whose knowledge requires management and containment, and another for those whose knowledge warrants amplification and infrastructure support.

An Epistemic Risk-Calibration Model governs communities' positions on the censorship-curation spectrum. Chronic, contested, or stigmatized conditions—where medical

knowledge remains ambiguous—gravitate toward restrictive censorship, raising legitimacy barriers precisely where experiential knowledge proves most critical. General health or acute conditions adopt integrative curation models. That calibration produces Diagnostic Determinism: legitimate participation depends on diagnosis, creating platform-mediated inequity. A user's ability to contribute meaningfully becomes preconditioned by their medical condition rather than their communicative competence or experiential authority. A Fragmented Toolscape—isolated, non-integrated governance tools—exacerbates cognitive load and navigational friction, disproportionately burdening low-literacy users. Four emergent governance clusters map a spectrum from credibility-through-restriction to inclusivity-through-curation, demonstrating that communities adopt coherent, systematic strategies rather than ad hoc responses to governance challenges.

Three theoretical contributions emerge from the analysis. First, to knowledge organization and platform governance studies, KO extends beyond static taxonomies; community rules and infrastructure function as dynamic, normative instruments of epistemic governance. The analysis moves beyond describing moderation outcomes to explaining the systemic production of epistemic inequality through sociotechnical systems, identifying core generative phenomena—Governance-Induced Epistemic Silencing, Infrastructural Amplification of Inequality, and the Participatory-Curatorial Divide—that explain how digital platforms institutionalize knowledge hierarchies.

Second, the integrated Epistemic Risk-Calibration Framework synthesizes these phenomena into a predictive, generalizable model. The framework posits governance posture as a function of perceived epistemic risk, generating specific infrastructural arrangements and social roles. That formulation effectively bridges KO theory, governance studies, and science and technology studies, providing a coherent explanation for variation in governance across knowledge communities. Communities do not adopt governance models arbitrarily but respond systematically to the epistemic characteristics of their knowledge domains.

Third, the findings advance theories of epistemic (in)justice in digital contexts by operationalizing how testimonial and hermeneutical injustice are technologically mediated. The systematic, rule-based invalidation of low-literacy contributions constitutes structural testimonial injustice, undermining credibility as knowers. Such invalidation is not incidental but embedded in the rule architecture itself. The Fragmented Toolscape and absence of scaffolding infrastructure exemplify technologized hermeneutical injustice, failing to provide collective interpretive resources that could bridge literacy gaps. Patients seeking support encounter systems that systematically devalue their experiential knowledge while privileging biomedical expertise, reproducing the very hierarchies that drove them to peer spaces.

Actionable implications follow for multiple stakeholders. For community designers and moderators, governance must move beyond default, risk-averse prohibitions to mitigate epistemic injustice. Active disruption of the Participatory-Curatorial Divide requires designing bridging infrastructure—templates that help users structure personal narratives to invite evidential connection, collaborative annotation tools allowing high-literacy users

to contextualize rather than segregate low-literacy posts, or structured feedback mechanisms that validate experiential knowledge without abandoning evidentiary standards. Moderators should engage in reflexive calibration, consciously using the framework to select a governance model suited to their community's context rather than unconsciously defaulting to containment for risky topics. For platform architects, the near-total absence of integrative and scaffolding tools highlights a critical design gap. Platforms should develop integrated knowledge ecosystems addressing the Fragmented Toolscape through built-in features for auto-linking to vetted resources, deeply integrated community wikis, and visual knowledge maps that reduce cognitive load and scaffold participation. Such features would reduce infrastructural amplification of inequality by making credible information more accessible to users regardless of literacy level.

For health communicators, policymakers, and researchers, the framework serves as a diagnostic and evaluative tool to audit community equity and design interventions. Communities serving the most vulnerable populations are architecturally predisposed to restrictive governance that may isolate those in greatest need. Partnerships and public health initiatives must therefore focus on co-designing risk-informed but equity-advancing governance features, ensuring efforts to disseminate credible information account for and actively reshape the underlying epistemic architecture determining who is heard and what counts as knowledge.

Several limitations contextualize these findings and point to future research directions. The analysis relies on declared, formal rules, offering a crucial view of codified norms but cannot capture the full spectrum of moderation-in-practice, including unwritten norms, moderator discretion, or user adaptation and resistance to rules. Ethnographic or interview-based studies are needed to understand the lived experience of such governance architecture. While the sample of 117 communities is substantive and diverse, it draws from a single large platform, shaping the kinds of governance tools observed and limiting generalizability to all online health spaces such as Facebook groups or dedicated forums. The Epistemic Risk-Calibration Framework requires validation and potential refinement through application in other knowledge domains—financial advice or climate forums—to test generalizability beyond health contexts. The cross-sectional, document-based design cannot trace governance evolution in response to specific crises or controversies, nor directly measure the causal impact of specific rules on user behavior, sense of belonging, or health outcomes. Longitudinal studies are needed to observe how communities dynamically shift along the KO-EQ spectrum. Finally, the analytical use of low-health-literacy and high-health-literacy ideal types, while necessary and revealing, necessarily simplifies a multifaceted literacy spectrum and may risk homogenizing diverse user motivations and competencies. Future research should combine rule analysis with behavioral data and user interviews to disentangle the complex interplay between literacy, identity, and the experience of epistemic (in)justice within these governed spaces.

7.0 Conclusion

The present study systematically explores how governance structures in online health communities navigate the core tension between credible knowledge and inclusive participation—the legitimacy dilemma. Using a mixed-methods approach analyzing 738 rule statements across 117 communities, the research shows that this dilemma is not just an abstract issue but is concretely embedded in a system of epistemic governance that actively shapes whose knowledge is valued and how it spreads. The findings reveal a consistent, measurable asymmetry: contributions aligned with high-health-literacy (HHL) norms are validated through credentialing and infrastructural support for curation, while those aligned with low-health-literacy (LHL) norms—mainly personal stories—are managed through Governance-Induced Epistemic Silencing via containment, channeling, and prohibition. This regulatory framework is reinforced by a Non-Neutral Infrastructure, in which tools such as structured formats and flair systems operationalize the epistemic boundary between "validated information" and "subjective experience," thereby widening literacy-based inequalities. These dynamics create a Participatory-Curatorial Divide, a social split that cements health literacy disparities into fixed roles within the community hierarchy. Importantly, this system is guided by an epistemic risk-calibration model: the way the dilemma is resolved is not random but a context-driven adjustment to the perceived epistemic risks within a community's disease domain. Forums dealing with chronic, disputed illnesses tend to adopt restrictive, censorship-heavy models, while those focused on general or infectious diseases lean toward more inclusive, curatorial approaches. This leads to a form of Diagnostic Determinism, where a user's right to epistemically legitimate participation depends on their diagnosis. The main theoretical contribution of this work is the development of the Epistemic Risk-Calibration Framework, which combines these phenomena to explain how digital platforms institutionally reproduce and technologize hierarchies of epistemic authority. By validating biomedical discourse while systematically marginalizing experiential knowledge, online health communities reproduce—through rules and infrastructure—the power dynamics that patients seek to bypass in traditional healthcare settings, thereby enacting structural testimonial and hermeneutical injustices in ostensibly peer-support environments. For practitioners, this analysis highlights the need to move beyond default governance models based on prohibition and risk mitigation. Creating equitable knowledge communities requires deliberate bridging tools—mechanisms and rules that support participation, incorporate diverse forms of evidence, and encourage collaborative curation rather than reinforcing a contributor-curator divide. Platform designers, moderators, and health policymakers must realize that balancing credible knowledge and inclusive support is possible without a zero-sum approach, but it demands intentional redesign of the socio-technical systems that mediate care and knowledge. Future research should extend this foundational mapping by examining moderation practices, the lived experiences of these governance structures, and their long-term effects. Additionally, the proposed framework invites testing and refinement across other contested domains—from climate science forums to financial advice communities—to develop a broader theory of epistemic governance in digital spaces. Ultimately, this study reveals the often-hidden mechanisms that shape digital health futures.

AI Usage Statement: Artificial intelligence (AI) tools were used solely for grammatical verification, language revision, and proofreading of the manuscript. AI was not used to generate the research findings, conduct the analysis, develop the methodology, or identify, generate, or verify references and citations. The authors remain fully responsible for the content, accuracy, interpretation, and conclusions presented in this work.

References

Berger, Peter L., and Thomas Luckmann. 1966. The Social Construction of Reality: A Treatise in the Sociology of Knowledge. Garden City, NY: Doubleday.

Dahlberg, Ingetraut. 1978. “A Referent-Oriented, Analytical Concept Theory for INTERCONCEPT.” International Classification 5: 142–151. https://doi.org/10.5771/0943-7444-1978-3-142.

Dourish, Paul. 2001. Where the Action Is: The Foundations of Embodied Interaction. Cambridge, MA: MIT Press.

Foucault, Michel. 1980. Power/Knowledge: Selected Interviews and Other Writings, 1972–1977. Edited by Colin Gordon. Translated by Colin Gordon, Leo Marshall, John Mepham, and Kate Soper. New York: Pantheon Books.

Fricker, Miranda. 2007. Epistemic Injustice: Power and the Ethics of Knowing. Oxford: Oxford University Press.

Gillespie, Tarleton. 2018. Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media. New Haven: Yale University Press.

Haraway, Donna. 1988. “Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective.” Feminist Studies 14: 575–599. https://doi.org/10.2307/3178066.

Hjørland, Birger. 2008. “What Is Knowledge Organization (KO)?” Knowledge Organization 35: 86–101. https://doi.org/10.5771/0943-7444-2008-2-3-86.

Nölke, Laura, Monika Mensing, Alexander Krämer, and Claudia Hornberg. 2015. “Sociodemographic and Health-(Care-)Related Characteristics of Online Health Information Seekers: A Cross-Sectional German Study.” BMC Public Health 15: 31. https://doi.org/10.1186/s12889-015-1423-0.

Norman, Cameron D., and Harvey A. Skinner. 2006. “eHealth Literacy: Essential Skills for Consumer Health in a Networked World.” Journal of Medical Internet Research 8, no. 2: e9. https://doi.org/10.2196/jmir.8.2.e9.

Nutbeam, Don. 2000. “Health Literacy as a Public Health Goal: A Challenge for Contemporary Health Education and Communication Strategies into the 21st Century.” Health Promotion International 15: 259–267. https://doi.org/10.1093/heapro/15.3.259.

Olson, Hope A. 2002. The Power to Name: Locating the Limits of Subject Representation in Libraries. Dordrecht: Kluwer Academic Publishers.

Szostak, Rick, Claudio Gnoli, and María López-Huertas. 2016. Interdisciplinary Knowledge Organization. Cham: Springer. https://doi.org/10.1007/978-3-319-30148-8.

Zarcadoolas, Christina, Andrew Pleasant, and David S. Greer. 2005. “Understanding Health Literacy: An Expanded Model.” Health Promotion International 20: 195–203. https://doi.org/10.1093/heapro/dah609.