The Knowledge Cosmos: A Playful, Immersive Platform for Exploring Science

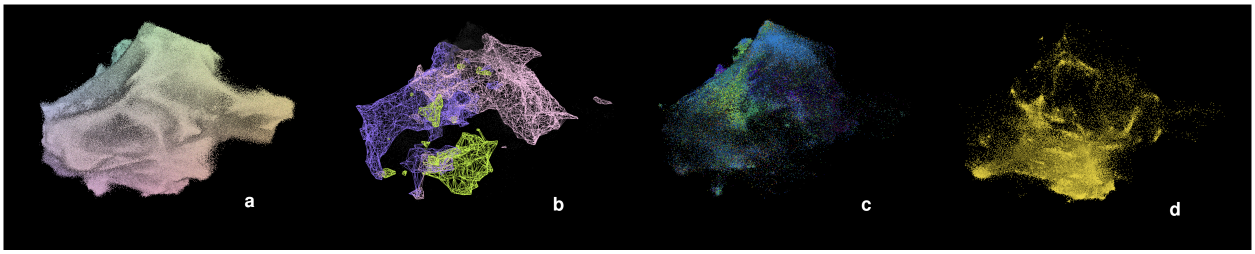


Figure 1: This image exhibits four views of The Knowledge Cosmos. From left to right: a. Full map of science, b. Constellations for Art, Medicine, and Machine Learning, c. Biology papers colored by subdiscipline, d. Papers at intersection of Biology and Medicine.

Abstract

As the volume of scientific literature continues to expand exponentially, traditional research tools struggle to keep pace—often reinforcing disciplinary silos and limiting opportunities for discovery. The Knowledge Cosmos re-imagines research exploration through an interactive, 3D visualization platform that treats science not as a static repository, but as a navigable universe. By spatializing 17 million academic papers based on semantic similarity, the platform enables users to explore the structure of knowledge intuitively, uncover interdisciplinary connections, and identify underexplored intellectual gaps.

Drawing on the principles of play, immersion, and serendipity, The Knowledge Cosmos democratizes the bird's eye view of research and encourages curiosity-driven inquiry across a wide range of users including students, educators, independent thinkers, and lifelong learners. This paper outlines the conceptual foundations, design, and technological infrastructure of the platform, shares insights from preliminary usability testing, and discusses future directions to scale its potential as a catalyst for interdisciplinary exploration and knowledge creation.

Index Terms: Large-scale data visualization, Science and technology visualization, User-centered design and evaluation, Knowledge visualization and concept mapping, Interdisciplinary research, 3D and spatial interaction, Visual encoding and design

1 Introduction

> *"Imagination is more important than knowledge, for knowledge is limited, while imagination encircles the world, stimulates progress, gives birth to evolution."* — Albert Einstein

The exponential growth of scientific literature presents a paradox: while access to information has never been greater, the ability to synthesize across disciplines and uncover novel insights has become increasingly difficult. Traditional research discovery tools, designed for targeted queries and linear retrieval, often reinforce disciplinary silos and discourage open-ended exploration. This raises a critical question: how might we make the vast landscape of research more accessible, navigable, and inclusive, especially for users who do not yet know what they are looking for?

The Knowledge Cosmos[1] addresses this challenge by reimagining research as a journey rather than a destination. It treats knowledge seeking not as an extractive act but as an act of care, visualizing 17 million academic papers in a 3D immersive space where proximity reflects semantic and disciplinary relationships. By spatializing science, the platform allows users to explore relationships visually, uncover unexpected intersections, and identify gaps across domains.

The project reflects collective care in three interconnected ways. It expresses care for knowledge by visualizing the architecture of intellectual effort: how disciplines relate, overlap, or remain siloed. It embodies care for curiosity by supporting creative wandering and open-ended discovery, encouraging users to ask questions before forming conclusions. And it advances care for the collective by revealing blind spots in academic systems and surfacing underexplored intersections that hold potential for more inclusive and impactful futures.

At the heart of the platform is play. Drawing on play as a core design principle, The Knowledge Cosmos reframes exploration as a joyful and intuitive process, more akin to stargazing than to traditional search. Users can zoom, pan, rotate, and filter through a vibrant celestial map of research, discovering intersections and patterns through dynamic visual feedback. Gamified features lower cognitive barriers and foster engagement across diverse audiences, including students, educators, artists, and independent learners. Beyond its immersive interface, The Knowledge Cosmos is designed to support serendipitous discovery. Users often encounter surprising connections—for example, between machine learning and biology or economics and environmental science—that would be difficult to surface through keyword-based search alone. By making the act of research feel exploratory, visual, and relational, the platform helps users form new connections, imagine novel directions, and participate in the collective evolution of knowledge.

The remainder of this paper is organized as follows. We begin with a literature review that situates The Knowledge Cosmos within broader conversations around creativity, interdisciplinarity, and the role of play in scientific discovery. We then turn to related work in the domain of research visualization tools, comparing existing platforms and approaches while identifying key limitations in how they support discovery, navigation, and user agency. Next, we present a detailed breakdown of our tool, the data architecture, and the de-

[1] Explore the tool at https://theknowledgecosmos.com/

sign principles that inform its interactive visual interface. We then share our approach and the results of initial user testing, including feedback on usability, engagement, and areas for improvement. We conclude with a discussion of future directions and the broader implications of a more open, collaborative, and navigable research landscape.

2 Interdisciplinarity and Exploration

Our tool is motivated by two interrelated forces that reverberate throughout the history of scientific discovery: interdisciplinarity and playful exploration.

Spectroscopy, initially a chemist's tool, became essential to decoding the light of distant stars. Computer science, once a purely technical field, now underpins our understanding of language, learning, evolution, and consciousness. Meanwhile, scientific visualization borrows from art and design to make the invisible both visible and meaningful. Such cross-disciplinary syntheses are a hallmark of high-impact science. Recent work in the "science of science" literature reinforces this: research that draws on a broader set of disciplines is more likely to be cited in patents, signaling greater technological relevance [25], and tends to score higher on disruption indicators, reflecting a greater likelihood of paradigm-shifting impact [6]. Recombining distant ideas, a practice often associated with interdisciplinary work, is especially vital in mature or stagnant fields, where incremental progress dominates [39]. One additional discipline in a research front is associated with a 20% boost in field-normalized citation impact [33].

Despite this growing empirical consensus, institutional frameworks still struggle to accommodate such work. Interdisciplinary research faces steeper barriers to publication [16], tends to appear in lower-impact journals [45], and is less likely to receive competitive funding [24]. The result is a structural disincentive to exactly the kind of inquiry most likely to yield transformative returns. These findings converge on a common theme: explorative research, as defined by March [30], is both high-risk and high-reward. These dynamics reveal not only the difficulty of interdisciplinary research but also the need for tools that support intellectual risk-taking, honor the complexity of disciplinary boundaries, and empower users to explore freely. The Knowledge Cosmos is designed with these values in mind, encouraging imaginative experimentation and supporting a broader spectrum of inquiry than traditional systems allow.

Playfulness, too, is a serious engine of insight. Richard Feynman famously approached physics like a playful tinkerer, cracking safes at Los Alamos, playing bongos, and freely experimenting with equations on his path to quantum electrodynamics. Claude Shannon, the father of information theory, filled his lab with unicycles, juggling machines, and flaming trumpets because for him, play was a form of problem-solving. Geneticist Barbara McClintock spoke of "a feeling for the organism," following intuitive hunches to discover jumping genes, for the fun of it [20]. These researchers exemplify an epistemic posture that is often critical for innovation. As Landers [23] reminds us, scientific reasoning is not reducible to logic or deduction; it also involves metaphor, recombination, and aesthetic sensibility. This matters for how we educate, and for how we build scientific tools. Well-designed gamified systems can improve not just learning outcomes, but also self-efficacy and engagement [22, 26, 44]. Kapp [18] draws a crucial distinction: gamification can make tedious tasks bearable, and more importantly, it can make valuable activities joyful. His iconic example: turning a staircase into a playable piano. The goal of ascending remains remains unchanged, but the experience is transformed. We believe science tools can, and should, work the same way.

The Knowledge Cosmos, too, is a playful hybrid — bridging human-computer interaction with scientometrics, cognitive science, and machine learning. It is not just a technical contribution, but a methodological and conceptual experiment in what becomes possible when we build for exploration.

3 Related Work

A wide range of efforts has emerged to map, visualize, and spatialize knowledge, enabling users to navigate the increasingly complex landscape of research. Foundational contributions, including Börner's Atlas of Science [4], established the groundwork for visualizing science at scale, while large-scale citation network analyses revealed the high-level organization of scientific fields for the first time [3, 32, 37].

Early spatializations of scientific knowledge translated abstract research domains into tangible, navigable forms but generally remained static and two-dimensional [12, 34, 43]. PaperScape [35] and arXiv Atlas [1] stand out, offering interactive maps of the arXiv corpus, enabling users to explore the relation and intersection of research fields and to follow citations around the map. Later systems introduced three-dimensional metaphors, using topographic landscapes where document density formed virtual "mountains." [9, 48]

Visualizing interdisciplinary intersections has become a central challenge in science mapping. Systems such as Galex [27], TopicPanorama [46], and PivotSlice [17] offered hierarchical and radial representations of overlapping research areas, primarily oriented toward expert users. More recently, Open Knowledge Maps [21] has lowered the barrier for public audiences by presenting topic clusters in simple, interactive formats.

Capturing the dynamics of scientific evolution has motivated a parallel stream of research. Systems such as CiteSpace [5], TextFlow [47], and CiteRivers [15] provided mechanisms for visualizing how research fields split, merge, and evolve over time. Later work introduced alluvial diagrams to track topic flows [38], while Memiescape [28] visualized multi-scale semantic evolution through "phylomemies."

Alongside academic systems, widely-used research discovery tools have advanced access to interdisciplinary literature. Google Scholar [14] enables broad, cross-domain search but offers primarily linear, text-based navigation. Connected Papers [8] and ResearchRabbit [36] introduce graph-based visualizations of related works, yet exploration typically remains confined to single-topic neighborhoods. Dimensions [10] supports multidisciplinary filtering and related research suggestions but largely follows structured query paths. Collectively, these platforms have expanded the means of accessing scientific knowledge; however, they often emphasize targeted retrieval over open-ended exploration, are primarily targeted at researchers, and none incorporate immersive, spatial metaphors for interdisciplinary discovery.

Building on the shoulders of these exceptional efforts, The Knowledge Cosmos offers a distinct contribution to the evolving landscape of research visualization. Departing from static maps, structured queries, and linear interfaces, it introduces an immersive, three-dimensional environment in which academic papers are rendered as spatially distributed points within a navigable universe. This dynamic spatialization invites users to explore the structure of scientific knowledge at multiple scales—revealing both local research neighborhoods and global disciplinary constellations.

The platform advances the identification of interdisciplinary intersections and thematic overlaps through features such as interactive visual filtering and field-based color encoding. Users can observe the diversity in the density of rendered clusters of papers to uncover areas of concentrated research activity and detect sparsely populated regions that may represent intellectual gaps or underexplored domains. Time-based filters further reveal how research areas evolve, offering a longitudinal view of knowledge formation.

By fusing semantic organization with spatial immersion, The Knowledge Cosmos reimagines research discovery as a journey that not only helps users navigate what is already known, but also illuminates where new knowledge might take shape.

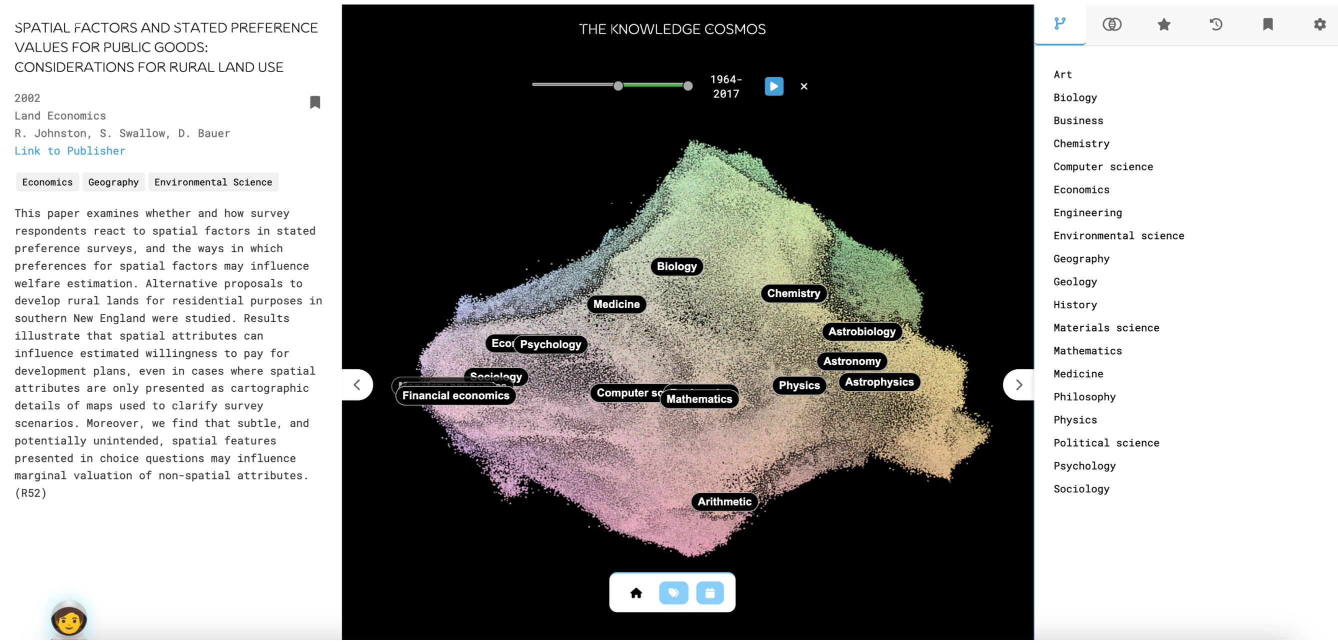


Figure 2: This image exhibits key features of the tool including the year filter, discipline labels, and selected paper panel on the left. Users can filter by top-level field using the right panel, explore intersections, constellations, and their history and saved papers.

4 Design

At the core of The Knowledge Cosmos is a dynamic 3D point cloud visualization, where each point represents a research paper (Figure 1a). Users can freely navigate this space using intuitive controls: rotating the view by dragging with the mouse, scrolling to zoom in and out, and moving through the cloud using keyboard controls (WASD/Arrow keys). When the user is outside the cloud looking in, dragging rotates the cloud. When the user is inside looking around, dragging rotates the camera.

When a user selects a point in the visualization, they are presented with comprehensive information about the corresponding paper, including the paper's title, abstract, authors, publication date, venue, a link to the publisher, and disciplinary tags. The history and saved papers tabs store what papers the user has viewed and saved during their journey, for future reference.

The Knowledge Cosmos organizes papers into a hierarchical structure of academic disciplines. In the *Subdisciplines* tab, users select a top-level field (e.g. Biology), to display color-coded biology papers by their subdiscipline (e.g. Biochemistry, Genetics, Ecology, etc.), see Figure 1c. This also expands a color legend in the Subdisciplines tab. By clicking on these legend items the user can filter further, showing only points from a specific subdiscipline.

The navigation toolbar is always present at the bottom of the map, offering a Home option to bring users back to the default view, a Labels option to show the names of disciplines on the map, and a Time filter that will filter to papers published in a specific range of years (see Figure 2). When the user navigates to the *Intersections* tab, the map is re-colored by top-level field. Users are asked to choose a top-level field (e.g. Computer Science), whereupon the map is filtered to only papers in that field, and re-colored based on the other top-level field each paper also belongs to (e.g. Art). By choosing a second field, the user can filter to only the points in that intersection (see Figure 1d). Our *Constellations* feature draws glowing wireframes which spatially map the boundaries of where the majority of papers in a discipline are. Constellations can be used to visualize the spatial relationships between disciplines, and to find intersecting spaces. Figure 1b uses constellations to compare the locations of Art, Medicine, and Machine Learning.

To support both new and experienced users, The Knowledge Cosmos offers an integrated tutorial system. First-time visitors can launch a guided tour through the "Take Me On A Tour" option, while ongoing assistance remains available at any time through Celeste, a virtual guide stationed in the lower-left corner of the interface. The tutorial introduces core features, navigation controls, and exploration strategies, ensuring users can quickly and confidently engage with the platform's full capabilities.

Every design decision, from the immersive interface to the interactive constellation overlays, aims to foster exploration without rigid boundaries or predetermined paths. The interface encourages users to linger, follow intuitive leads, and uncover meaningful relationships that might otherwise remain invisible. Features like guided tours and a virtual assistant lower the barrier for first-time users, ensuring the experience is welcoming and empowering regardless of disciplinary background or technical fluency. This emphasis on openness and ease of access reflects a broader goal: to create a space where inquiry feels supported, and where a wide range of perspectives can engage with the evolving research landscape.

5 Application and Insights

The Knowledge Cosmos empowers users to move beyond keyword-based search and instead explore research through spatial, semantic, and disciplinary relationships. By interacting with its key features—disciplinary filtering, constellation overlays, intersection mapping, and temporal sliders—users can surface novel insights, form cross-domain connections, and identify both dense research clusters and sparsely populated areas. Below, we demonstrate applications of these features and the types of discoveries they enable.

5.1 Surfacing Underexplored Intersections

By filtering for research that spans Geography and Medicine, users can locate papers that address global health challenges through a spatial lens. One such paper, "Effective Coverage and Systems Effectiveness for Malaria Case Management in Sub-Saharan African Countries" [13], see Figure 3a, appears in a relatively sparse cluster.

This scarcity suggests that while global health is a rich field, the systems-level intersection between geography, public health systems, and medical intervention remains underexplored. Such gaps offer promising areas for further inquiry, particularly in light of climate change and shifting disease ecologies. Similarly, filtering for Business and History reveals another sparse intersection. A standout paper, ”Consumer Perceptions and Willingness to Pay for Intrinsically Motivated Online Content” [29], see Figure 3b, integrates theories from management, digital platforms, and information systems history. However, few surrounding papers indicate that the historical evolution of digital economies, particularly the ways business models, user values, and content monetization intersect, remains relatively underexplored. This presents a valuable opportunity for scholars in media studies, economics, and digital humanities to examine longitudinal trends in digital consumption, reputation systems, or platform capitalism. By drawing attention to these sparse yet promising intersections, The Knowledge Cosmos acts as a lens for locating new intellectual terrain, encouraging users not only to follow existing paths, but also to forge new ones.

5.2 Revealing Dense Interdisciplinary Clusters

While sparsely connected disciplines can highlight research gaps, densely populated intersections reveal areas of intense academic activity—often where disciplinary boundaries are blurred by shared methods, problems, or frameworks. When filtering for papers that span Biology and Medicine, see Figure 4, users uncover a thick, richly connected region in the map, indicating a longstanding and active exchange between the two fields. This cluster spans topics such as molecular biology, pharmacology, epidemiology, and biomedical engineering.

What makes this feature especially powerful is the ability to visualize subdisciplinary diversity within the cluster. Users can see, for instance, how genomics research flows into personalized medicine, or how microbiological studies inform drug resistance strategies. The high-density structure of this intersection confirms the strength of the relationship in addition to inviting more nuanced exploration, prompting queries such as: Where are the densest thematic cores? Which areas are slightly peripheral, hinting at under-tapped niches like synthetic biology or systems immunology? By navigating these dense clusters, users can trace intellectual lineages, uncover overlapping citations, and find interdisciplinary work that may otherwise be siloed by journal or discipline. This spatial overview allows for more strategic literature reviews and research planning, helping users find central ideas as well as overlooked edges within well-trodden domains.

5.3 Navigating Constellations to Understand Disciplinary Interconnectedness

The constellation overlay feature visualizes how disciplines cluster and connect, offering an intuitive sense of their semantic and epistemic proximity. In one view (Figure 5), clusters related to Art, Machine Learning, and Medicine, form contiguous regions, suggesting emerging interdisciplinary fields such as AI-assisted medical imaging or computational creativity. The visual adjacency of these constellations opens up speculative lines of inquiry: How might artists and machine learning experts collaborate on tools for diagnostics? What ethical frameworks emerge at these intersections? By contrast, disciplines like Art and Environmental Science may appear more distant, yet the spatial separation itself can spark curiosity: what conceptual bridges might be built to connect them? These spatial relationships can serve as a visual grammar of academic interconnectedness. Users can use this feature to identify well-integrated fields or disciplinary silos that may benefit from more dialogue. Exploring these distances invites users to not only observe the current state of knowledge, but also imagine new, meaningful proximities.

5.4 Exploring Knowledge Evolution Over Time

The time filter in The Knowledge Cosmos allows users to dynamically trace how interdisciplinary relationships or siloed disciplines develop, intensify, or transform. Unlike static bibliographic tools, it reveals not only when connections first appear but also how quickly they proliferate. A particularly striking example is the intersection between Computer Science and Biology. When filtered for publications spanning 1900 to 2000, the intersection appears relatively sparse, with a modest distribution of research scattered across the map. However, extending the filter just 17 years forward to 1900–2017 reveals a dramatic densification (Figure 6). Entire regions of the map become saturated with points, reflecting an explosive growth in work that blends computational methods with biological inquiry. This surge corresponds to the rise of fields like bioinformatics, computational genomics, and systems biology, areas that have become central to modern science yet barely registered on the research landscape just two decades earlier. What was once a loosely connected set of papers has matured into a robust interdisciplinary cluster, suggesting both increased collaboration and the emergence of shared frameworks and tools.

5.5 Unique Formations When Mapping Science

One feature which immediately stands out to the user is the shape of science as projected in 3D space. It appears as a spiny napkin-ring topology. There are numerous great and small crests which jut from the galaxy, one for mathematical physics, another for agroforestry, fisheries and paleontology, another holding our knowledge of veterinary medicine, to name a few. A thin line of papers connects abstract algebra with theoretical physics, sporting a continuum of theories of the nature of the universe. Only once the user becomes more familiar with the galaxy would they notice the two entrances, one on top and another on the bottom, which give entry to the hollow interior. See Figure 7 for some views of the variegated spatial structure of scientific knowledge.

6 METHODS

The scientific papers we visualize are sourced from a subset of Microsoft Academic Graph [42], giving significant coverage of academic works and citations between them. The positions for papers are computed from the approximately 17 million semantic embeddings provided by Ke et al. [19], so-called SPECTER embeddings [7]. Papers are close in this 768-dimensional space to the extent that their titles and abstracts are semantically similar. To produce these the authors begin with embeddings from the SciBERT [2] model, which has been trained to predict missing words in scientific text and whether one sentence follows another. These embeddings are then fine-tuned using the citation links between papers.

We use UMAP [31] to embed these high-dimensional vectors into 3D space. UMAP preserves both local and cluster-level distances, while remaining computationally efficient, making the projection of these 17 million papers achievable on consumer hardware. UMAP's closest alternative, t-SNE, is much more computationally intensive, but also can distort global topology, potentially placing fields which are quite distant in the high-dimensional space close together in three dimensions.

To visualize these papers' 3D positions, we modify Potree [40], an open-source WebGL renderer originally developed for LiDAR point clouds. Potree dynamically adjusts point density based on distance, enabling seamless rendering at massive scales. Paper positions and IDs are bundled in a LiDAR binary format and converted to Potree's octree format. When the user clicks on a paper, the interface queries the Semantic Scholar API [41] for further information - citation links, field names, extended abstract, and authors - on demand. We also store publication year in the .las file as the source_id, which is range-filterable in Potree. For each top-level academic field, we create two additional point clouds, one for subdisciplines and another for

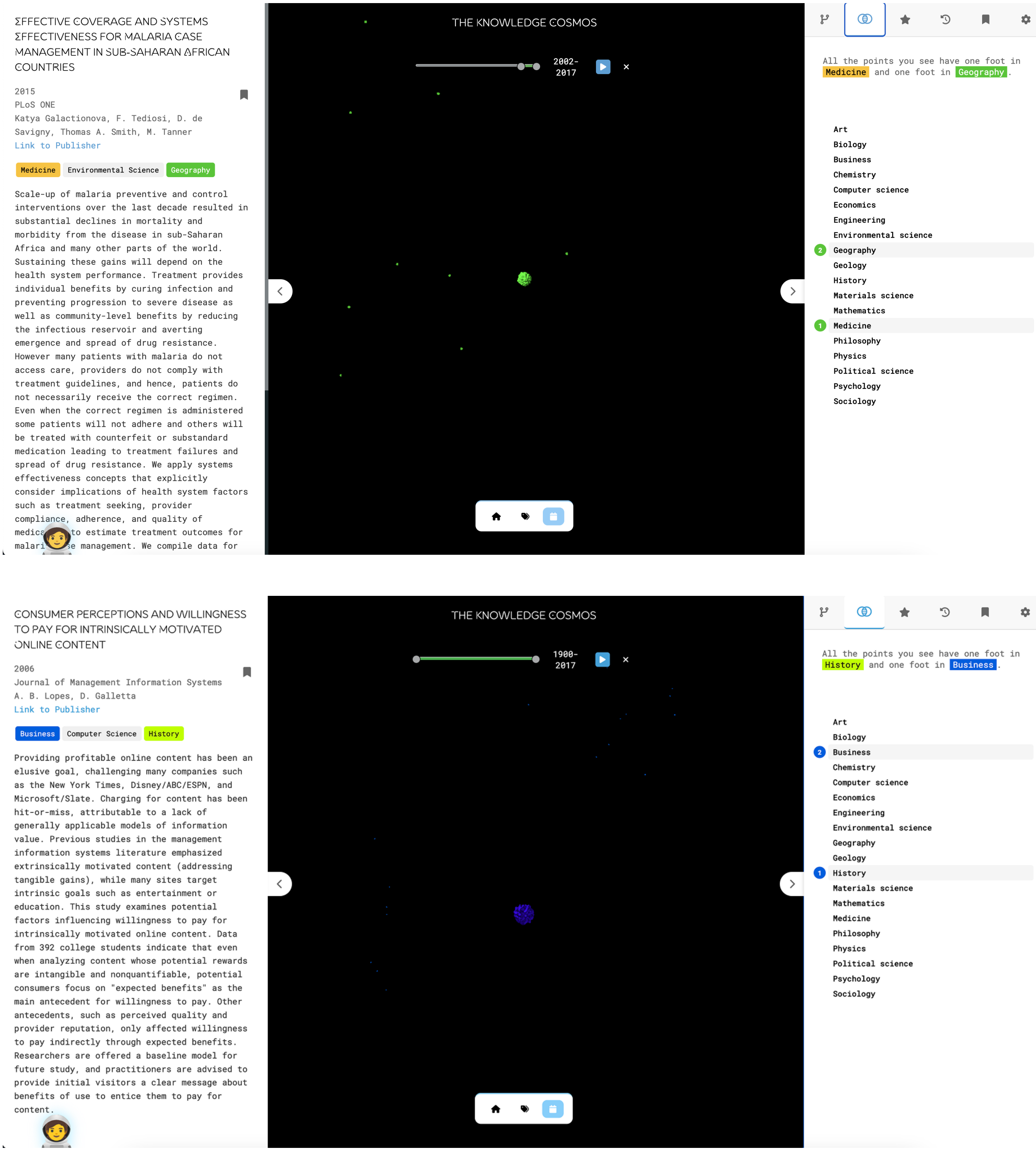


Figure 3: The cluster of papers intersecting (a) Geography and Medicine, top, and (b) Business and History, bottom, are relatively sparse, offering promising areas for further inquiry.

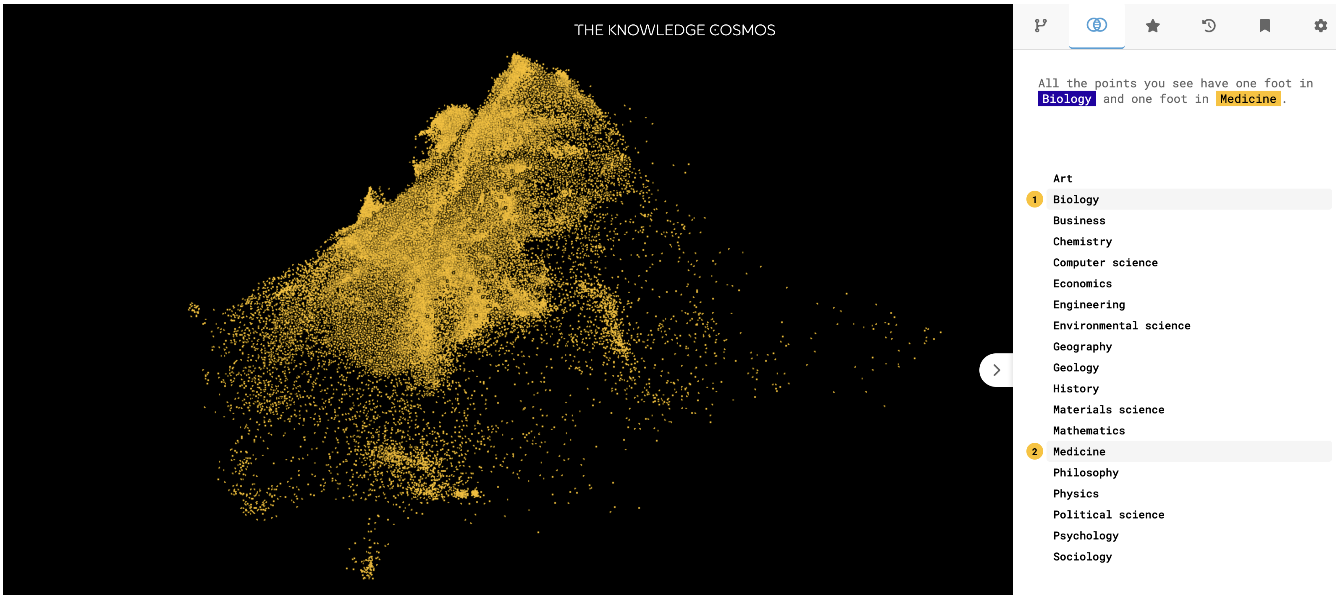


Figure 4: The density of paper intersecting Medicine and Biology is dense, suggesting a high degree of distinct subdisciplines, reflecting specialized domains such as genomics, immunology, pharmacology, and neuroscience.

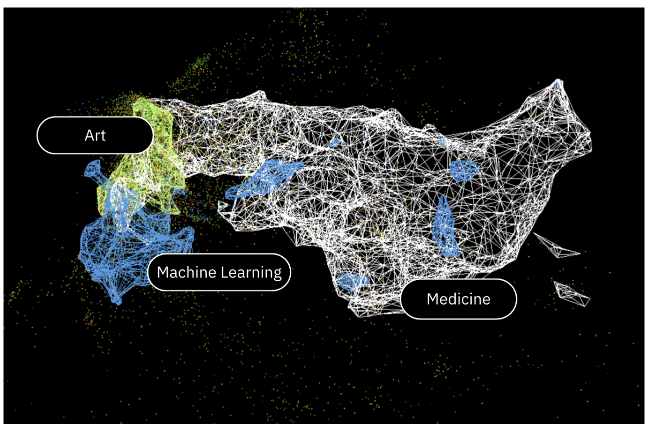


Figure 5: The constellations for art, machine learning, and medicine, show the boundaries of each field on the map.

intersecting fields. In these we store the subdiscipline or intersection for each paper as an integer classification, which allows dynamic coloring and filtering.

Constellations are 3D polyhedra which show the approximate boundaries of academic fields. We compute these polyhedra using the α-shape algorithm [11]. Intuitively, given a fixed α, the algorithm draws faces between three points whenever a ball of radius $1/\alpha$ can be drawn such that the boundary contains the three points, and no points are contained inside. Before computing the α-shape, we exclude points which are below a fixed local density, ensuring constellations show the core of each field and not its total extent. These 3D polyhedra are rendered as wireframes, so as to not obscure the points in the cloud. When any constellation is visible, point cloud opacity is reduced for clarity.

Labels for fields are positioned at each field's constellation's center of mass, adjusted to the nearest vertex if this center is outside the polyhedron. Clicking a label reveals the corresponding constellation, and reorients the camera to a position outside the cloud looking through the constellation at the center of the cloud.

The webpage is designed for static hosting, allowing for an inexpensive and scalable architecture hosted on AWS S3, and requiring no database software. All 3D graphics are rendered client-side using THREE.js, supported by WebGL, primarily through shaders provided by Potree.

7 Usability Testing and Discussion

To examine the preliminary version of The Knowledge Cosmos, we conducted a usability evaluation and tested the tool's efficacy in supporting interdisciplinary research exploration. We recruited 6 participants with varied backgrounds, including students, researchers, and general knowledge seekers. Participants represented a range of familiarity with data visualization tools and research discovery platforms to ensure diverse perspectives.

Participants completed five structured tasks designed to evaluate core functionalities, with data collected on objective success, task completion time, and self-reported difficulty. Performance varied notably across activities. Response times and subjective ease are shown in Figure 8. Task 3 (Finding Intersections) emerged as the most challenging, exhibiting the longest average completion time (M = 3 minutes 45 seconds, SD = 1 minute 45 seconds) and the lowest subjective ease ratings. This finding informed the addition and design of the *Intersections* tab in our tool today. In contrast, Task 2 (Applying a Constellation) was completed most efficiently, with mean completion times under one minute and consistently high success ratings. Aggregate difficulty ratings indicated that Tasks 1 (Applying a Filter), 2 (Applying a Constellation), and 4 (Saving and Retrieving a Paper) were generally well understood (mean difficulty scores between 2.0 and 2.5 on a 3-point scale). Across all tasks, completion rates were 100%. The overall usability score, based on a 5-point Likert scale, averaged 3.25 (SD = 0.4), suggesting moderate ease of use with meaningful areas for refinement.

Qualitative feedback further illuminated both strengths and targeted opportunities for improvement. Participants consistently praised the visual and aesthetic appeal of the 3D cosmic interface, describing the experience as playful, engaging, and markedly more enjoyable than conventional research tools. This novelty sparked a strong desire for continued exploration and serendipitous discov-

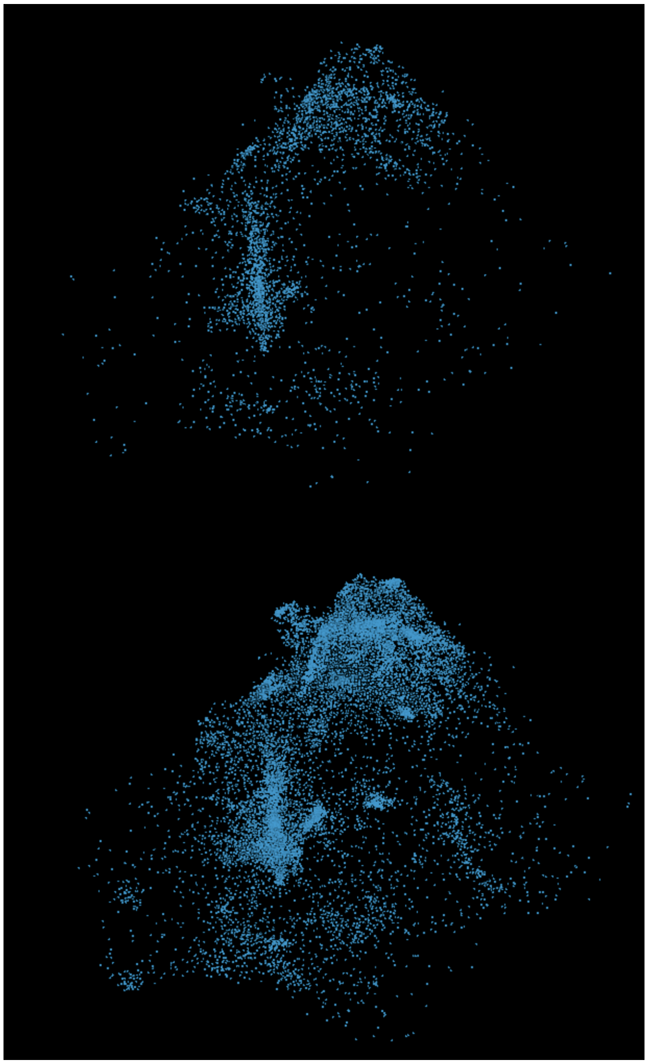


Figure 6: Evolution of papers intersecting Computer Science and Biology from 1900–2000 (top) and 1900–2017 (bottom).

ery, aligned with the platform's mission to support interdisciplinary research. The feedback also reinforces the value of designing for curiosity, enabling users to follow questions they didn't know they had, and to discover insights through open-ended engagement. Several participants also commented on unexpected intersections between fields, indicating that the platform not only informs but also broadens conceptual horizons.Task 5 (freeform exploration) exhibited the second longest completion time, indicating a strong user inclination toward continued engagement with the tool.

Several opportunities were also identified. Participants noted inconsistencies in disciplinary categorization; for example, expected subdisciplines such as Comparative Politics were either missing or placed under unexpected hierarchies, disrupting intuitive exploration. Difficulties with spatial navigation—including sensitivity of movement controls and disorientation within the 3D space—were common, suggesting a need for improved calibration. The limitation of the dataset to papers published before 2017 was also noted, particularly by participants exploring rapidly evolving domains like Machine Learning and Art. Additionally, participants expressed interest in advanced features for curating saved papers (e.g., creating thematic lists or saving exploration paths) and recommended embedding guidance systems or starting points to assist newcomers in navigating the research landscape. These findings underscore the value of inclusive, participatory design in knowledge tools, where users shape their own learning journeys and contribute to ongoing refinement.

Given the small sample size (n=6), inferential statistics were not applied; rather, findings were triangulated through descriptive statistics and thematic qualitative analysis. The results collectively surface critical interaction and interface design challenges that will guide future iterative development. Additional iterative testing with a larger user base is necessary to validate recent design decisions and better understand emerging user needs.

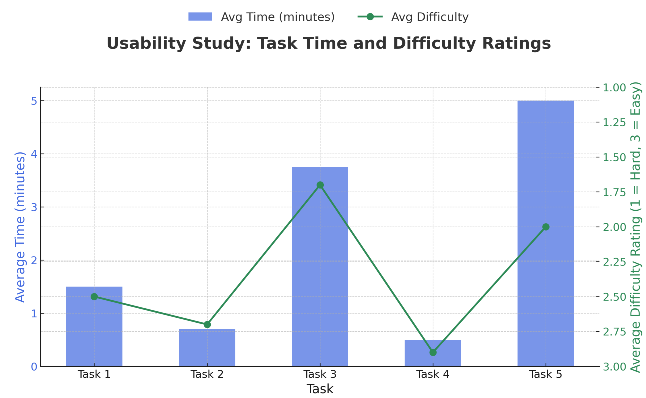


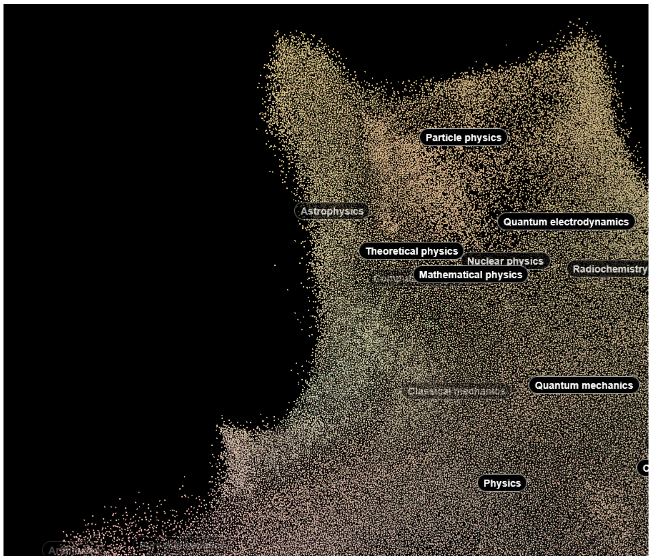
Figure 8: Average task completion time (bars) and user-reported difficulty ratings (line) for each usability task.

8 Conclusion and Future Work

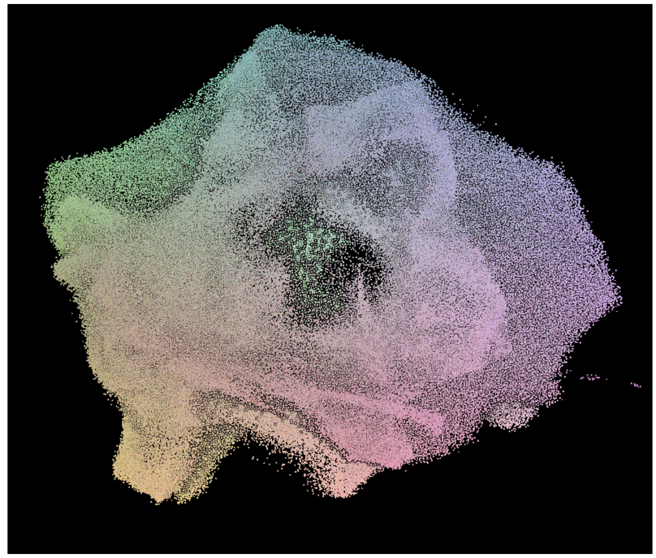
The Knowledge Cosmos presents a first step toward reimagining how its users explore, connect, and contribute to the vast universe of human knowledge. Usability testing validated that The Knowledge Cosmos actively fosters curiosity and sparks novel lines of inquiry with participants describing the experience as immersive and exciting. The tool's intuitive design lowered the threshold for engagement, empowering users to pursue their interests fluidly across disciplinary boundaries. Importantly, even non-expert users engaged deeply with the content, such as one participant who explored intersections between her own field of Political Science and her peer's background in Computer Science as a way to bridge their domains.

Usability testing also illuminated several directions for future development which could significantly enhance the platform's capabilities. Currently, the platform is based on a dataset of approximately 17 million research papers. While substantial, this represents only a fraction of the available global research corpus. Expanding the dataset to encompass a broader, more up-to-date collection — ideally scaling to the more than 350 million indexed journal articles, and continually indexing papers as they are published — would help ensure relevance across disciplines and timeframes. Expanding the dataset beyond traditional journal articles would further strengthen the platform's utility, particularly for disciplines such as Art and Anthropology, which often produce and communicate knowledge through mediums such as ethnographic texts or visual works.

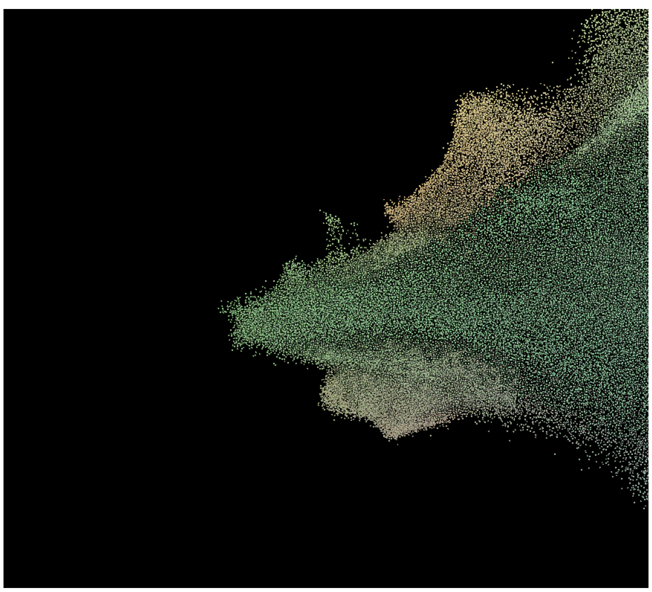
In addition, rigorous accessibility testing will be imperative to ensure the platform is inclusive and usable across a wide range of devices, abilities, and contexts. Participants expressed a desire for features such as hover-based previews, as well as the ability to search by author or keyword. There is also significant potential to use bibliometric techniques and artificial intelligence to guide users. For instance, algorithms could surface pivotal papers, uncover unexpected intersections between disciplines, and identify emerging gaps or opportunities for research.



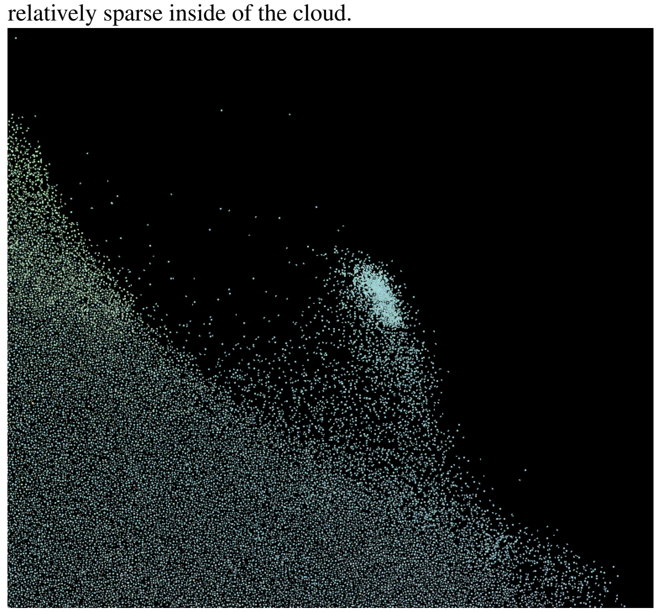
(a) An imposing crest of astrophysics and related fields fills the right of the figure, bridging through the center and down to the left into abstract mathematics. The small node which is seen jutting from the lower-middle of this figure connects most closely with engineering, and involves the design of electrical systems for spacecraft and aerospace.



(b) When viewing The Knowledge Cosmos from below the user can glimpse that it is hollow - that there is a massive hole in the middle which cuts straight through the bottom. Traveling along the ring counter clockwise the user would find computer vision, speech recognition, audiology, physical medicine and rehabilitation, orthodontics, biomedical engineering, electrobiology, and medical imagine physics. The top of the cloud also admits an entrance to the relatively sparse inside of the cloud.



(c) The shape of this crest is intriguing, with multiple subordinate crests jutting from it. Each structure represents a different subfield of chemistry, including organometallic, catalysis and reaction engineering, and application spaces such as sustainable energy. The sparse whisp visible rising from the crest, in the center of the image, consists of a few hundred crystal diffraction studies.



(d) This odd structure seems to consist solely of studies in the biophysics of protein aggregation, particularly amyloid aggregation, which is relevant to neurodegenerative diseases like Alzheimer's and Huntington's.

Figure 7: Stills of the shape of scientific knowledge.

As the dataset grows in size and complexity, optimizing the platform's underlying architecture will be critical to maintaining smooth and responsive interactions. Ensuring scalability, performance, and user experience at larger scales will be central to The Knowledge Cosmos' evolution into a powerful and enduring research tool.

In the long term, The Knowledge Cosmos seeks to be a space where knowledge is not merely consumed, but tended to, cultivated, and shared. As the dataset grows and accessibility expands across platforms, our aim is to support a more diverse constellation of learners, thinkers, and communities. Future development will focus on deeper personalization and more intuitive guidance, so that users not only find information, but also form meaningful, self-directed connections across disciplines. Ultimately, we envision the platform contributing to a more generative knowledge ecosystem, where discovery is not framed solely as retrieval, and scholarship is a shared endeavor.

References

[1] arXiv Atlas. https://github.com/Jaluus/ArXivAtlas. 2

[2] I. Beltagy, K. Lo, and A. Cohan. SciBERT: A Pretrained Language Model for Scientific Text, Sept. 2019. arXiv:1903.10676 [cs]. doi: 10.48550/arXiv.1903.10676 4

[3] K. W. Boyack, R. Klavans, and K. Börner. Mapping the backbone of science. *Scientometrics*, 64(3):351–374, Aug. 2005. doi: 10.1007/s11192-005-0255-6 2

[4] K. Börner. *Atlas of Science: Visualizing What We Know*. The MIT Press, Boston, MA, 2010. 2

[5] C. Chen. CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature. *Journal of the American Society for Information Science and Technology*, 57(3):359–377, Feb. 2006. doi: 10.1002/asi.20317 2

[6] D. Cheng, Z. Xue, Y. Zhibo, and Z. Mingze. Impact of interdisciplinarity on disruptive innovation: the moderating role of collaboration pattern and collaboration size. *Scientometrics*, Mar. 2025. doi: 10.1007/s11192-025-05285-3 2

[7] A. Cohan, S. Feldman, I. Beltagy, D. Downey, and D. Weld. SPECTER: Document-level Representation Learning using Citation-informed Transformers. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pp. 2270–2282. Association for Computational Linguistics, Online, 2020. doi: 10.18653/v1/2020.acl-main.207 4

[8] Connected Papers. https://www.connectedpapers.com/. 2

[9] G. S. Davidson, B. Hendrickson, D. K. Johnson, C. E. Meyers, and B. N. Wylie. Knowledge Mining With VxInsight: Discovery Through Interaction. 2

[10] Dimensions.ai. https://www.dimensions.ai/. 2

[11] H. Edelsbrunner, D. Kirkpatrick, and R. Seidel. On the shape of a set of points in the plane. *IEEE Transactions on Information Theory*, 29(4):551–559, July 1983. doi: 10.1109/TIT.1983.1056714 6

[12] D. Fried and S. G. Kobourov. Maps of Computer Science. In *2014 IEEE Pacific Visualization Symposium*, pp. 113–120. IEEE, Yokohama, Mar. 2014. doi: 10.1109/PacificVis.2014.47 2

[13] K. Galactionova, F. Tediosi, D. De Savigny, T. Smith, and M. Tanner. Effective Coverage and Systems Effectiveness for Malaria Case Management in Sub-Saharan African Countries. *PLOS ONE*, 10(5):e0127818, May 2015. doi: 10.1371/journal.pone.0127818 3

[14] Google Scholar. https://scholar.google.com/. 2

[15] F. Heimerl, Q. Han, S. Koch, and T. Ertl. CiteRivers: Visual Analytics of Citation Patterns. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):190–199, Jan. 2016. doi: 10.1109/TVCG.2015.2467621 2

[16] S. M. Hyrynsalmi, G. Liebel, R. d. S. Santos, and S. Baltes. Not real or too soft? On the challenges of publishing interdisciplinary software engineering research, Jan. 2025. arXiv:2501.06523 [cs]. doi: 10.48550/arXiv.2501.06523 2

[17] Jian Zhao, C. Collins, F. Chevalier, and R. Balakrishnan. Interactive Exploration of Implicit and Explicit Relations in Faceted Datasets. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2080–2089, Dec. 2013. doi: 10.1109/TVCG.2013.167 2

[18] K. M. Kapp. *The gamification of learning and instruction: game-based methods and strategies for training and education*. Pfeiffer, San Francisco, CA, 2012. 2

[19] Q. Ke, L. Liang, Y. Ding, S. V. David, and D. E. Acuna. A dataset of mentorship in bioscience with semantic and demographic estimations. *Scientific Data*, 9(1):467, Aug. 2022. doi: 10.1038/s41597-022-01578-x 4

[20] E. F. Keller. *A Feeling for the Organism: The Life and Work of Barbara McClintock*. W.H. Freeman, 1983. 2

[21] P. Kraker and N. Shaghaei. Open Knowledge Maps: a Visual Interface to the World's Scientific Knowledge. June 2019. Publisher: Zenodo. doi: 10.5281/ZENODO.3258105 2

[22] G. Lampropoulos and Kinshuk. Virtual reality and gamification in education: a systematic review. *Educational technology research and development*, 72(3):1691–1785, June 2024. doi: 10.1007/s11423-024-10351-3 2

[23] A. L. Landers. Ordinary Creativity in Patent Law: The Artist within the Scientist. *Missouri Law Review*, 75, 2010. 2

[24] G. Laudel. The art of getting funded: how scientists adapt to their funding conditions. *Science and Public Policy*, 33(7):489–504, Aug. 2006. doi: 10.3152/147154306781778777 2

[25] B. Li, S. Chen, and V. Larivière. Interdisciplinarity affects the technological impact of scientific research. *Scientometrics*, 128(12):6527–6559, Dec. 2023. doi: 10.1007/s11192-023-04846-8 2

[26] M. Li, S. Ma, and Y. Shi. Examining the effectiveness of gamification as a tool promoting teaching and learning in educational settings: a meta-analysis. *Frontiers in Psychology*, 14:1253549, Oct. 2023. doi: 10.3389/fpsyg.2023.1253549 2

[27] Z. Li, C. Zhang, S. Jia, and J. Zhang. Galex: Exploring the Evolution and Intersection of Disciplines. *IEEE Transactions on Visualization and Computer Graphics*, pp. 1–1, 2019. doi: 10.1109/TVCG.2019.2934667 2

[28] Q. Lobbé, A. Delanoë, and D. Chavalarias. Exploring, browsing and interacting with multi-scale structures of knowledge. 2

[29] A. B. Lopes and D. F. Galletta. Consumer Perceptions and Willingness to Pay for Intrinsically Motivated Online Content. *Journal of Management Information Systems*, 23(2):203–231, Oct. 2006. doi: 10.2753/MIS0742-1222230209 4

[30] J. G. March. Exploration and Exploitation in Organizational Learning. *Organization Science*, 2(1):71–87, Feb. 1991. doi: 10.1287/orsc.2.1.71 2

[31] L. McInnes, J. Healy, and J. Melville. UMAP: Uniform manifold approximation and projection for dimension reduction. *arXiv*, 2018. arXiv: 1802.03426. 4

[32] J. Moody and R. Light. A view from above: The evolving sociological landscape. *American Sociologist*, 37(2):67–86, 2006. doi: 10.1007/s12108-006-1006-8 2

[33] K. Okamura. Interdisciplinarity revisited: evidence for research impact and dynamism. *Palgrave Communications*, 5(1):141, Nov. 2019. doi: 10.1057/s41599-019-0352-4 2

[34] Pak Chung Wong, B. Hetzler, C. Posse, M. Whiting, S. Havre, N. Cramer, Anuj Shah, M. Singhal, A. Turner, and J. Thomas. IN-SPIRE InfoVis 2004 Contest Entry. In *IEEE Symposium on Information Visualization*, pp. r2–r2. IEEE, Austin, TX, USA, 2004. doi: 10.1109/INFVIS.2004.37 2

[35] PaperScape. https://paperscape.org/. 2

[36] ResearchRabbit. https://www.researchrabbit.ai/. 2

[37] M. Rosvall and C. T. Bergstrom. Maps of random walks on complex networks reveal community structure. *Proceedings of the National Academy of Sciences*, 105(4):1118–1123, 2008. doi: 10.1073/pnas.0706851105 2

[38] M. Rosvall and C. T. Bergstrom. Mapping Change in Large Networks. *PLoS ONE*, 5(1):e8694, Jan. 2010. doi: 10.1371/journal.pone.0008694 2

[39] A. Rzhetsky, J. G. Foster, I. T. Foster, and J. A. Evans. Choosing experiments to accelerate collective discovery. *Proceedings of the National Academy of Sciences of the United States of America*, 112(47):14569–14574, 2015. ISBN: 1509757112. doi: 10.1073/pnas.1509757112 2

[40] M. Schütz, S. Ohrhallinger, and M. Wimmer. Fast Out-of-Core Octree

Generation for Massive Point Clouds. *Computer Graphics Forum*, 39(7):155–167, Oct. 2020. doi: 10.1111/cgf.14134 4

[41] Semantic Scholar. https://www.semanticscholar.org/. 4

[42] A. Sinha, Z. Shen, Y. Song, H. Ma, D. Eide, B.-J. P. Hsu, and K. Wang. An Overview of Microsoft Academic Service (MAS) and Applications. In *Proceedings of the 24th International Conference on World Wide Web*, pp. 243–246. ACM, Florence Italy, May 2015. doi: 10.1145/2740908.2742839 4

[43] A. Skupin. The world of geography: Visualizing a knowledge domain with cartographic means. *Proceedings of the National Academy of Sciences*, 101(suppl_1):5274–5278, Apr. 2004. doi: 10.1073/pnas.0307654100 2

[44] L. Smirani and H. Yamani. Analysing the Impact of Gamification Techniques on Enhancing Learner Engagement, Motivation, and Knowledge Retention: A Structural Equation Modelling Approach. *Electronic Journal of e-Learning*, 22(9):111–124, Nov. 2024. doi: 10.34190/ejel.22.9.3563 2

[45] J. Wang, R. Veugelers, and P. Stephan. Bias against novelty in science: A cautionary tale for users of bibliometric indicators. *Research Policy*, 46(8):1416–1436, Oct. 2017. doi: 10.1016/j.respol.2017.06.006 2

[46] X. Wang, S. Liu, J. Liu, J. Chen, J. Zhu, and B. Guo. TopicPanorama: A Full Picture of Relevant Topics. *IEEE Transactions on Visualization and Computer Graphics*, 22(12):2508–2521, Dec. 2016. doi: 10.1109/TVCG.2016.2515592 2

[47] Weiwei Cui, Shixia Liu, Li Tan, Conglei Shi, Yangqiu Song, Zekai Gao, Huamin Qu, and Xin Tong. TextFlow: Towards Better Understanding of Evolving Topics in Text. *IEEE Transactions on Visualization and Computer Graphics*, 17(12):2412–2421, Dec. 2011. doi: 10.1109/TVCG.2011.239 2

[48] J. Wise, J. Thomas, K. Pennock, D. Lantrip, M. Pottier, A. Schur, and V. Crow. Visualizing the non-visual: spatial analysis and interaction with information from text documents. In *Proceedings of Visualization 1995 Conference*, pp. 51–58,. IEEE Comput. Soc. Press, Atlanta, GA, USA, 1995. doi: 10.1109/INFVIS.1995.528686 2