

From Top-Down to Organic Evolving Graphs, Ontologies, and Taxonomies

by Dinis Cruz and ChatGPT Deep Research, 2025/03/29

Introduction

The way we organize knowledge is undergoing a paradigm shift. Traditional top-down approaches to ontologies and taxonomies — where experts rigidly define schemas in advance — are giving way to more **organic, adaptive models** that evolve over time. Instead of imposing a fixed hierarchy from above, modern knowledge graphs and ontologies are often shaped dynamically by data and community input. Researchers and practitioners now argue that schemas should **grow and refine themselves continuously**, much like a living system, rather than remain static blueprints.

This research document examines why a dynamic approach is superior, looking at theoretical frameworks and real-world successes that champion **bottom-up evolution** over top-down design. Key themes include the pitfalls of rigid knowledge modeling, the benefits of continuous refinement, the crucial role of human-in-the-loop oversight, and how generative AI (Large Language Models) can turbocharge ontology development.

We also highlight examples like Wikipedia's sprawling category network and collaborative cybersecurity knowledge bases to illustrate how flexible taxonomies thrive in practice.

Finally, we consider "ontologies of ontologies" as a meta-layer of abstraction and stress the importance of intuitive design and visualization in making these evolving graphs usable.

Pitfalls of Top-Down Knowledge Modeling

Top-down knowledge modeling has historically been the default — experts define a canonical taxonomy or ontology, and all data must fit into that predetermined structure. **This approach has notable shortcomings:**

- **Inflexible Hierarchies:** Rigid top-down taxonomies often force information into strict tree structures. Real knowledge, however, rarely fits neatly into a single hierarchy. For example, a traditional taxonomy can struggle with concepts that belong in multiple categories (polyhierarchy). One industry analysis notes that **strict parent-child hierarchies are limiting**; adding an extra dimension (like a country spanning two continents) causes ambiguity when only “broader/narrower” relations are allowed ([Why a Knowledge Graph is the Best Way to Upgrade Your Taxonomy - Enterprise Knowledge](#)). In practice, rigid schemas can't gracefully handle such nuances.
- **Incomplete & Outdated Schemas:** No team of designers can anticipate all concepts and relationships in a complex domain upfront. Domains like cybersecurity or biology change rapidly, and a top-down ontology can become outdated quickly. Manually building and maintaining ontologies is also **resource-intensive and slow**, leading to the notorious “knowledge acquisition bottleneck” ([Ontology Learning Process as a Bottom-up Strategy for Building Domain-specific Ontology from Legal Texts](#)). In other words, by the time a committee updates the schema, reality may have moved on.
- **Lack of Expressivity:** Simpler taxonomies (e.g. basic hierarchies) capture only limited relationships. They usually allow just “is-a” or part-whole links, missing rich contextual connections. This constrains the ontology's usefulness. As one knowledge engineering firm points out, a pure taxonomy “**cannot infer or recommend**” new insights because it lacks the relational depth of a graph; advanced applications like NLP or recommendations are “not possible without enhancing your taxonomy with an ontology”. Thus, sticking to a spartan schema limits intelligent use of the knowledge.
- **Poor Alignment with Users' Mental Models:** Top-down designs may reflect how experts think the domain *should* be organized, rather than how practitioners or data naturally cluster information. This misalignment can hurt usability. Clay Shirky famously argued that many classic categorization schemes are a “**bad fit**” for the messy reality of the electronic world; imposing rigid order on information that users tag and link in organic ways will inevitably clash with actual usage ([Shirky: Ontology is Overrated – Categories, Links, and Tags](#)). In short, authoritative schemas can become brittle or irrelevant if they ignore the more fluid, folk categorization that real users prefer.

Given these pitfalls, it's clear why **strict standardization often falters**. Even well-designed ontologies (for example, the early Semantic Web vocabularies) faced adoption issues when they were too static or complex for broad community buy-in. In practice, some ambitious top-down efforts ended up as “paper ontologies” with little live use. These lessons set the stage for a more flexible, evolutionary approach to knowledge modeling.

Benefits of Organic Evolution in Graph-Based Structures

In contrast to the static top-down method, an **organic, evolving approach** to ontologies and graphs treats the schema as a living construct that adapts with usage and new information. There are several key benefits to this dynamic paradigm:

- **Continuous Refinement:** An evolving knowledge graph can be updated incrementally as new concepts emerge or relationships change. Instead of large, infrequent overhauls, the schema undergoes **constant small improvements**. This means the model stays current with the domain. In cybersecurity, for instance, threat tactics are continuously evolving, and the community-driven ATT&CK framework stays relevant by being in a “*constant state of development*” – always refining and extending to cover new attacker behaviors ([MITRE ATT&CK®](#)). The ability to tweak the ontology in near real-time is invaluable in rapidly changing fields.
- **Bottom-Up Emergence of Structure:** When users or data contributions drive the schema, the ontology can capture how information *naturally* organizes itself. Patterns that were not obvious to the original designers can surface. This bottom-up emergence often produces a **more intuitive and comprehensive structure** than a top-down guess. In the early days of Web 2.0, user tagging systems (folksonomies) demonstrated that “free-form labeling, without regard to categorical constraints” could nonetheless yield useful organization from the crowd’s collective efforts. Big, messy data sets tagged by users revealed logical groupings on their own. In other words, order can emerge from apparent chaos, aligning closely with how people actually think about the content.
- **Greater Flexibility and Contextualization:** An organic graph isn’t limited to one rigid hierarchy; it can accommodate multiple classification axes and nuanced relationships. New relationship types or attributes can be added as needed, providing **rich context** for each entity. A more flexible ontology might specify custom relations (e.g. “is partially in [X]”) or add metadata that a simple taxonomy would forbid. This flexibility prevents the model from “boxing in” knowledge. Each addition makes the graph smarter and more descriptive, rather than breaking the schema.
- **Community Engagement and Ownership:** When the ontology evolves through contributions (with appropriate oversight), stakeholders feel a sense of ownership and engagement. A living knowledge model that incorporates user feedback and domain expert input can rally a community

of practice around maintaining it. This **human element** often leads to higher quality and resilience: errors get corrected, and the ontology better reflects consensus reality. We see this in open projects like Wikipedia, where thousands of editors organically adjust categories and links. The result is “an **organically evolving category system** that reflects the current needs of the user-contributor community” ([Visualizing large-scale human collaboration in Wikipedia](#)). In such cases, the taxonomy’s evolution is guided by many eyes, making it robust and relevant.

In summary, an organically evolving graph or ontology is *adaptive*. It grows like a healthy ecosystem, continually adjusting and optimizing. This stands in sharp contrast to ossified top-down designs that risk obsolescence. By embracing constant evolution, organizations can maintain **living knowledge bases** that stay aligned with reality and user needs.

Human-in-the-Loop: Marrying Automation with Curation

A critical ingredient in successful organic knowledge models is the **human-in-the-loop**. While automation can propose new schema elements or relationships (especially with AI advances we’ll discuss later), human expertise is vital to validate and refine these structures. In practice, the best approach is *not* purely bottom-up or purely top-down, but a **continuous collaboration between machine suggestions and human judgment**:

- **Semi-Automatic Ontology Evolution:** Researchers have developed frameworks for **ontology learning**, where algorithms extract candidate concepts and relations from data (e.g. text corpora) and suggest them for inclusion in the ontology. These act as bottom-up feeds to grow the knowledge graph. However, human oversight is usually built in to ensure quality. As one study explains, ontology learning (OL) is typically “a **semi-automatic process** where the ontology engineer and the domain expert can be involved to achieve better results,” with their expertise used to verify information and decide what is valuable (). In other words, machines do the heavy lifting of finding patterns, but humans curate the results to ensure the schema makes sense.
- **Validation and Refinement:** Human domain experts serve as **editors and validators** for an evolving ontology. They can confirm if a proposed new class or relationship truly belongs, or if a reclassification is warranted. This role is similar to Wikipedia editors monitoring changes: the crowd may suggest a new category or linkage, but experienced editors (or consensus discussions) validate it. A human-in-the-loop process guards against the ontology devolving into nonsense or overly fragmented categories. It provides a feedback mechanism so that automated additions improve the graph’s quality rather than degrade it.

- **Continuous User Feedback:** Beyond formal ontology engineers, everyday end-users can implicitly shape the schema through their interactions. For example, in a semantic search system, if users consistently fail to find something under the current taxonomy, that signals a needed change. A dynamic system can solicit **user feedback loops** (ratings, corrections, usage analytics) to inform how the ontology should adapt. The human perspective ensures that the knowledge model remains **practical and intuitive**, not just theoretically complete.
- **Governance vs. Emergence Balance:** The challenge is finding the sweet spot between letting the ontology grow freely and enforcing quality control. Some governance – guidelines for contributions, periodic reviews – is necessary to maintain coherence. Projects like **MITRE's ATT&CK** for cybersecurity have a clear process: they welcome community contributions and new technique proposals to expand the framework, but changes are vetted and released in curated updates (roughly every 6 months). This combination of open contribution with editorial control exemplifies human-in-the-loop governance that keeps an evolving knowledge base credible.

In essence, **human stewardship is what turns a merely evolving graph into a *reliably* evolving graph**. The loop of machine generation and human curation leverages the best of both: scale and speed from algorithms, plus wisdom and contextual understanding from people. Together, they ensure the semantic network improves over time in a valid, trusted way.

Generative AI and LLMs as Ontology Catalysts

One of the most exciting developments enabling organic schema evolution is the rise of **Generative AI and Large Language Models (LLMs)**. These AI systems can analyze vast amounts of text or data and help **automatically infer ontologies or mappings** at a scale previously impossible. Rather than replacing human experts, LLMs augment them by handling tedious ontology engineering tasks and surfacing patterns. Here's how generative AI is supporting scalable, contextual ontology creation:

- **Rapid Ontology Bootstrapping:** Traditionally, designing an ontology from scratch was a manual, consultant-driven process that could take months. LLMs can dramatically accelerate this by ingesting domain literature and proposing an initial set of entities, categories, and relationships. Recent work shows that **LLMs excel at capturing semantic relationships** and even hierarchical structures from natural language ([Knowledge Graph\(s\) and LLM-based ontologies have a very good shot at unlocking GenAI in production | by EQT Ventures | eqtventures | Medium](#)). They effectively distill the implicit ontology that is latent in unstructured text. For instance, given a corpus of medical research, an LLM

might outline a draft ontology of diseases, symptoms, treatments, and their interrelations. This provides a running start, which humans can then refine.

- **Continuous Ontology Updates:** Because LLMs are built on transformer models that can be retrained or prompted with new data, they are **easy to update** with the latest knowledge. This is crucial for maintaining an evolving knowledge graph. Instead of manually adding new concepts, one could periodically feed the LLM new articles or records and have it suggest updates to the ontology. The AI essentially learns the changes in the domain and reflects them in the schema. This dynamic updating aligns with governance needs: as fields change, the AI keeps the ontology current, and humans oversee the integration of those changes.
- **Ontology Alignment and Mapping:** Large Language Models can also serve as translators or mediators between different knowledge schemas. In large organizations or across the open web, multiple ontologies often need to be reconciled (e.g. mapping one institution's taxonomy to an industry standard). LLMs, with their broad exposure to language, can understand that, say, "*CRM Customer*" in one database likely corresponds to "*Client*" in another. They can generate candidate mappings or even a **meta-ontology that links equivalent or related concepts** across systems. This capability helps realize the idea of a **contextual meta-taxonomy**, where instead of forcing everyone into one schema, an ontology-of-ontologies can bridge between many. Generative AI provides the scalability to maintain these complex mappings.
- **Inferring Rules and Constraints:** Beyond just class and relation suggestions, advanced generative models can propose **logical rules or constraints** that give an ontology more structure. For example, after reading documentation, an AI might infer that "if X is a subtype of Y, then every X must have property Z" and suggest adding that as an ontological rule. Recent analyses note that LLMs can assist not only in defining ontology concepts but also in defining **symbolic rules** that govern the knowledge graph's reasoning. This helps the evolving graph maintain consistency and derive new facts (inferences) as it grows.

By turbocharging the creation and refinement of ontologies, AI ensures that our knowledge graphs can scale without collapsing under manual labor. One venture analysis put it succinctly: until now, ontology design was a manual bottleneck done by consultants, but "**LLMs are a powerful way to build ontologies, enabling knowledge graphs to be built and updated much faster**". The role of AI is thus that of catalyst and assistant curator, working hand-in-hand with human experts. This synergy allows for **truly dynamic semantic networks** that remain both comprehensive and context-sensitive as they evolve.

Case Studies: Evolving Taxonomies in the Wild

The theory of organic ontologies is compelling, but how does it play out in practice? Several real-world knowledge systems have embraced **flexible, evolving taxonomies** and achieved remarkable success. Here we highlight a few examples across different domains:

- **Wikipedia's Category Network:** Perhaps the most famous case of bottom-up classification at scale is Wikipedia. Every Wikipedia article can be tagged with multiple categories, and users can create new categories or reorganize them at any time. Over the years, this has produced an enormous, decentralized classification graph rather than a strict tree. Studies of Wikipedia's category system describe it as an **"organically evolving category system" reflecting the needs of its contributors**. Unlike a library classification devised by a central authority, Wikipedia's taxonomy grows and adapts as new topics arise. This openness does lead to varying granularity (some parts of the category graph are very fine-grained, others coarse), but it also means the structure can flexibly accommodate anything the encyclopedia covers. The success of Wikipedia — with millions of articles organized reasonably well without a top-down schema — showcases the power of a collaborative, ever-adjusting ontology. What emerges is a rich, tangled hierarchy of categories linked in complex ways, arguably an "ontology of the sum of human knowledge" that no single designer could have anticipated.
- **Wikidata and Collaborative Knowledge Graphs:** Sister to Wikipedia, **Wikidata** is a collaborative knowledge graph that anyone can edit. It provides a structured database of entities (people, places, concepts, etc.) and their relationships, used by Wikipedia and many other projects. Wikidata's ontology is not fixed; new properties and classes can be proposed by the community. This has allowed it to **cover an immense breadth of domains** in just a decade, from genealogy to astronomical objects, all under one evolving graph. The community-driven model lets niche experts extend the ontology in their area (e.g. specialists adding a new property for a specific scientific concept) while others validate it. As a result, Wikidata functions as a **living ontology** for the world, continually refined through thousands of daily edits. It stands as a real-world proof that decentralized ontology building can produce a high-quality, large-scale knowledge graph. In fact, Wikidata is often cited as **"a free collaborative knowledge graph"** underpinning many AI and data integration applications ([Can Knowledge Graphs Reduce Hallucinations in LLMs? : A Survey](#)).
- **MITRE ATT&CK in Cybersecurity:** The MITRE ATT&CK framework is a knowledge base of adversary tactics, techniques, and procedures used in cybersecurity. Rather than a one-time taxonomy, ATT&CK was designed to be **constantly updated** as threats evolve. MITRE actively solicits contributions from the security community — researchers and analysts can submit new attack techniques or suggest modifications to existing ones. These inputs are reviewed and, if accepted, incorporated into the framework's periodic releases. Since its public debut in 2015, ATT&CK has grown significantly, expanding into multiple domains (enterprise IT, mobile, cloud, ICS) with new categories and tactics added regularly ([How Has MITRE ATT&CK Improved, Adapted, and Evolved?](#)) ([Mitre ATT&CK Framework Simplified: Understanding Cyber Threats](#)). The community-driven evolution has been crucial to its success: it remains *the* reference for threat behavior because it keeps pace with reality. This is a stark contrast to

older cybersecurity taxonomies that became stale as new attack methods bypassed the old categories. ATT&CK's model of an **open, curated, evolving ontology** has since inspired similar efforts in related areas of cyber threat intelligence.

- **Enterprise Knowledge Graphs & Folksonomies:** In corporate settings, we also see recognition that taxonomies must be flexible. E-commerce giant Amazon, for example, uses a product taxonomy that is constantly adjusted based on inventory and user navigation patterns. In the early days of social media, folksonomy tag systems (like Delicious for bookmarks or Flickr for photos) showed that letting users organically tag content could outperform rigid classifications. As Shirky observed, individual tagging choices made for personal use aggregated into **“high group value”** organization when viewed collectively ([Clay Shirky - "Ontology is overrated": a review - Nic Höning](#)). Many modern enterprise knowledge management systems now combine a base taxonomy with user-generated tags and machine learning suggestions – essentially adopting a hybrid approach that ensures the information architecture can **adapt to how users actually label and search for content**.

Across these examples, a common thread is the **balance of freedom and oversight**. Wikipedia and Wikidata have policies and editors; ATT&CK has the MITRE team; enterprise systems have information architects – but none of these impose immovable rules that never change. Instead, their role is to guide the evolution, not to freeze it. The result in each case is a **resilient, richly interconnected knowledge structure** that stays relevant precisely because it was allowed to evolve.

Ontologies of Ontologies: A Meta Layer of Abstraction

As knowledge ecosystems mature, an interesting phenomenon occurs: we develop **ontologies of ontologies** – essentially, meta-taxonomies that organize multiple schemas. This is a natural layer of abstraction that emerges when different groups create their own ontologies or classification systems and we then need to **coordinate and integrate across them**.

Consider the biomedical domain: there are ontologies for genes, proteins, diseases, phenotypes, etc., built by different expert communities. Initiatives like the **Open Biomedical Ontologies (OBO) consortium** arose to provide a unifying layer. OBO serves as a sort of **“registry” of vocabularies** across biology, aiming to standardize how they interrelate (). In other words, OBO functions as an ontology of ontologies – a top-level structure to which all member ontologies adhere. This meta-ontology approach lets researchers fetch controlled vocabularies for various subdomains in a consistent way. It's effectively a *standard of standards*, ensuring that, for example, a gene ontology and an anatomy ontology can work together without collisions.

In the business world, the same concept appears when companies establish an **enterprise metadata schema** that maps together taxonomies from different departments. Academically, this idea is sometimes formalized: “A meta-ontology is a schema for other ontologies... a simple ontology whose concepts are super-classes for a further refined domain ontology.” ((PDF) [Building a business domain meta-ontology for information pre-processing](#)). In practice, that means defining an upper layer with very general concepts that individual domain ontologies specialize. For instance, a meta-ontology might define that any domain ontology will have concepts like *Entity*, *Attribute*, *Relation*, which domain-specific ontologies like “Customer ontology” or “Product ontology” then extend. This provides a consistent abstraction across all knowledge models in an organization.

Crucially, these meta-level frameworks should themselves be handled organically, not rigidly. We’re essentially **applying the organic principle at a higher layer**: just as domain ontologies must evolve, the integration schema (the ontology of ontologies) should also adapt as new domains or standards come into play. If one tries to fix the meta-ontology in stone, it could hinder the growth of the lower-level ontologies. Instead, the meta level should capture common patterns and allow alignment, while remaining extensible.

The concept of ontologies of ontologies highlights our growing sophistication in knowledge management. It acknowledges that no single ontology will rule them all; instead, we’ll have an **ecosystem of interlinked ontologies**, and we need abstract layers and mappings to organize that ecosystem. Embracing this view is part of moving beyond one-size-fits-all top-down models toward a **federation of evolving schemas** operating in a coordinated fashion.

Usability and Visualization of Dynamic Graphs

No matter how powerful a dynamically evolving knowledge graph or ontology is, it only delivers value if people can understand and use it. **Intuitive design and visualization** are paramount in making dynamic graphs accessible. When a schema is constantly changing, good UX becomes even more critical to prevent confusion. Key considerations include:

- **Clear Visual Representation:** Humans are visual creatures. Representing a knowledge graph in a clear, navigable format helps users grasp complex relationships quickly. **Knowledge graph visualization** techniques depict nodes and edges in an interactive graph, often with clustering or color-coding to convey different types of relationships. Done well, this provides “a clear and intuitive means of understanding and exploring intricate networks of data points” ([Knowledge graph visualization: A comprehensive guide \[with examples\]](#)). For example, an employee looking at

an evolving skills ontology could use a graph viewer to see how a certain skill connects to roles, training materials, and experts in the company. A visual interface lets them follow paths through the graph naturally, turning abstract data into a navigable map.

- **Dynamic Updates with Context:** In a changing ontology, the interface should highlight what's new or modified. Users need to trust the system, so showing versioning, annotations, or change history can help. Visualization tools can indicate which parts of the graph are **recently added or have shifting connections**, so users aren't blindsided by silent changes. Additionally, contextual information (like tooltips or sidebars with definitions) ensures that as new schema elements appear, their meaning is readily available. An intuitive design will bridge the gap between the *technical complexity* of an evolving semantic model and the *mental model* of the user.
- **Support for Multiple Perspectives:** A dynamic knowledge base often contains different facets and perspectives. Good design allows users to slice and filter the graph in ways that make sense to them. For instance, one might toggle between a hierarchical tree view (for a traditional outline of the taxonomy) and a graph view (to see cross-connections). Providing multiple visualization modes caters to both those who think in hierarchies and those who think in networks. It also helps illustrate why a dynamic graph is richer than a static tree — the user can literally see the additional relationships.
- **Lowering the Barrier to Entry:** Perhaps most importantly, intuitive design democratizes the ontology. A well-designed interface can enable even non-technical users to query the knowledge graph or contribute to it. Consider how Wikipedia made editing simple for laypeople; similarly, a dynamic ontology platform should make tasks like adding a link or suggesting a new category as straightforward as possible (with guardrails). When visualizations and UI are user-friendly, the **semantic network becomes a tool for everyday users**, not just ontologists. This increased participation further fuels the organic evolution, creating a virtuous cycle.

In sum, **visualization and UX are the translation layer** that turns an evolving, high-dimensional graph into human-understandable knowledge. As one guide notes, graph visualizations can “simplify the presentation of intricate concepts and relationships, making it easier for non-experts to grasp the content”. This clarity builds confidence and adoption. After all, the most elegantly adaptive ontology means little if stakeholders can't comprehend or navigate it. Investing in intuitive design is thus an integral part of implementing dynamic knowledge systems at scale.

Conclusion: Toward Living Knowledge Ecosystems

The future of knowledge engineering lies in **living, breathing ontologies and graphs** that grow organically rather than being chiseled in stone. Top-down designs served a purpose in an earlier era, but their limitations are increasingly evident in the face of exponential information growth and change. A new paradigm is emerging — one that combines *the wisdom of crowds, the prowess of AI, and thoughtful human curation* to build adaptive semantic structures.

By allowing graphs, ontologies, and taxonomies to evolve, we gain resilience and agility. Mistakes in the schema can be discovered and corrected, gaps filled, and new knowledge assimilated continuously. The process becomes one of ongoing learning, not one-time modeling. Human-in-the-loop workflows ensure that this evolution maintains quality and relevance, preventing chaos while embracing flexibility. Meanwhile, generative AI and LLMs act as force-multipliers, scanning the horizons of data to surface patterns and suggestions that keep the ontology up-to-date and context-rich.

We stand, therefore, on the cusp of **truly intelligent knowledge ecosystems**. These are systems that not only store information but also *adapt to it*, reorganize themselves, and even abstract themselves (via meta-ontologies) as they scale. Real-world successes from open communities and enterprises alike show that this approach is not just theoretical but practical and effective. The role of designers shifts from crafting perfect taxonomies to facilitating growth — setting initial conditions, then guiding and pruning the knowledge garden as it flourishes.

For organizations and practitioners, the mandate is clear: **embrace organic ontology evolution as a strategy**. Encourage collaborative schema development, invest in tools that support dynamic graphs, and leverage AI to handle complexity. Cultivate an “ontology of ontologies” mindset to integrate diverse knowledge sources. And always keep the end-user in focus through intuitive design, so that the sophisticated underpinnings translate into actual insights and value.

In doing so, we move closer to knowledge management that mirrors the way human understanding develops — not by static definitions, but by continuous refinement. The vision is a world where our knowledge graphs are as alive as the knowledge itself, constantly learning, adapting, and enriching their structure. This organic evolution of graphs, ontologies, and taxonomies promises a more robust and context-aware foundation for everything from enterprise data to scientific research to AI applications. It is an evolution in how we represent knowledge, driven by the very knowledge it represents, and it holds the key to unlocking smarter, more responsive information systems in the years ahead.