Digital Identity Optimization (DIO) and the Ontology of Digital Identity (ODI): A Multidimensional Analysis of Conceptual Coherence, Machine-Readable Representation, and Network Distribution

Author-Led Research Using DeepSeek

Daniel Beránek

Independent Researcher, Brno, Czech Republic

dan@danielberanek.cz

October 30, 2026

---

1. Executive Summary

This report examines the implementation and distribution of Daniel Beránek's Digital Identity Optimization (DIO) and Ontology of Digital Identity (ODI) through the framework's own conceptual logic, complemented by a graph-theoretical analysis. The central question is whether, at the observed point in time, the distributed DIO/ODI ecosystem appears to sustain a recognizable, referentially continuous, and meaningfully reconstructable identity across heterogeneous representations and entry points, despite partial inconsistencies, gaps, inaccessible resources, and uneven updating.

The central finding is that **within the inspected pathways, the DIO/ODI ecosystem does sustain a recognizable and meaningfully reconstructable identity**. The core conceptual structure—the distinction between DIO as applied discipline and ODI as theoretical meta-layer, the six dimensions of the semantic connectome (Entity, Representation, Interpretation, Trust, Relationship, Reconstruction), and the Reconstruction Failure Modes taxonomy—is recognizably continuous across the reference vocabulary, the Zenodo publication record, the manifesto corpus, the `llms.txt` context file, and the application mappings. The ecosystem provides multiple pathways to this structure, and the most consequential distinctions are restated across independent artifact types and genres.

The mechanisms supporting reconstruction are architectural rather than incidental. Stable concept URIs under a canonical namespace provide a referential anchor. The Zenodo corpus encodes provenance relations (`IsDerivedFrom`, `IsIdenticalTo`, `References`) that support temporal reconstruction by making version history inspectable. The Reference Model v0.5.0 explicitly describes itself as operating above the distributed corpus, consolidating and cross-referencing concepts developed across the Manifest, the Conceptual Framework, the *Crisis* paper, MVDIO, and the Taxonomic Note. The manifesto corpus declares the Czech version as the interpretively referential source while treating other language projections as coordinated but autonomous. The `llms.txt` file makes explicit the evidential distinctions and version boundaries that the narrative pages leave implicit, though these distinctions are also recoverable from other artifacts.

The strains are real but, within the inspected pathways, **locally contained rather than systemically compromising**. Multilingual hub-dependence concentrates interpretive authority in the Czech source and the central hub, leaving individual language projections as peripheral spokes; the core concepts remain recoverable, but the architecture creates the structural condition for semantic fragmentation if a revision propagates unevenly. Version misalignment across archived releases creates conditions in which an observer entering through an older record may encounter a different conceptual state than

one entering through the current Reference Model; the pathway to reconciliation exists through the consolidation node, but whether an observer traverses it depends on their entry point and diligence. The single-operator dependency is a structural vulnerability rather than an observed failure: the current artifacts are maintained and mutually consistent, but their continued coherence depends on maintenance concentrated in one person. The unresolved Wikidata notability question indicates that reconstruction within the open knowledge-graph infrastructure is not fully secured, though the author-entity remains stabilized by ORCID and Zenodo.

The `llms.txt` file is interpretively convenient but not structurally necessary. The evidential distinctions it makes explicit—professional biography versus research publication, live web text versus archived edition, canonical vocabulary versus derived application mapping—are also recoverable from the narrative corpus, the Reference Model's self-description, and the publication metadata. If the file were unavailable, an observer would lose a compressed, prioritized interface, but not the underlying conceptual structure.

The assessment is neither an unconditional endorsement nor a mechanical compliance audit. The framework's own concepts—referential anchoring, provenance relations, nodal points, contextual adequacy—provide a differentiated account of why the ecosystem holds together, where continuity is strained, and which weaknesses are material to reconstruction at different levels of detail. The weaknesses are locally contained within the observed snapshot; they create conditions under which reconstruction could fail for a particular observer entering through a particular route, but the ecosystem provides multiple pathways to the core conceptual structure, and the most consequential distinctions are restated across independent artifact types.

---

2. Research Approach & Corpus

2.1 Entry Sequence and Orientation

The research began with the machine-readable layer, as the framework's own architecture suggests. The **DIO/ODI Reference Vocabulary** (`vocab.danielberanek.cz/dio-odi/`) was inspected first, revealing a SKOS-first vocabulary (v0.2.0) with stable concept URIs under a canonical namespace (`https://vocab.danielberanek.cz/dio-odi/#`), broader/narrower and related-concept relations, provenance notes, and explicit serializations. The **GitHub repository** (`dan-beranek/dio-odi-reference-model`) provided access to the release history and the Mermaid concept-map source. The repository documentation explains the SKOS-first rationale: "A future OWL layer should add stronger axioms only where the published corpus and subsequent review justify them." This sequence established conceptual orientation before encountering the rhetorical density of the manifesto.

The `llms.txt` and `llms-full.txt` files were then analyzed as machine-readable context artifacts. `llms.txt` is a structured self-description that defines the author's identity, education, E-E-A-T categories, key pages, service architecture, research scope, reference artifacts, crawling priorities, and AI usage policy. `llms-full.txt` is a selected full-text export of eight source pages, with explicit export metadata stating the snapshot date (2026-10-29), the eight source documents, and the transformation applied (diacritics removed, ASCII-only, reversible Unicode notation for non-Latin scripts). These files constitute a declared machine-readable interface to the website's conceptual content.

The **Zenodo layer** was then explored. The author search page was initially inaccessible due to maintenance, but individual records were successfully accessed via direct DOI resolution and search. The following records were retrieved and inspected:

Record	Version	Date	DOI	Role
Manifest of DIO and ODI: Canonical Multilingual Edition	1.0.0	2026-07-26	10.5281/zenodo.21610934	Canonical dataset; archival root
DIO/ODI Reference Model	v0.3.1	2026-10-11	10.5281/zenodo.23290681	Early reference model
DIO/ODI Reference Model	v0.5.0	2026-10-25	10.5281/zenodo.23643345	Current consolidated model
Minimum Viable DIO (MVDIO)	1.0.0	2026-10-08	10.5281/zenodo.23226205	Analytical reduction
Taxonomic Note on Reconstruction Failure Modes	—	2026-10-09	10.5281/zenodo.23269123	Taxonomy revision
The Crisis of Entity Resolution in Heterogeneous AI Environments	1.0.0	2026-09-02	10.5281/zenodo.22257504	Problem statement
Conceptual Framework for Entity-Based Visibility	—	2026-08-07	10.5281/zenodo.21839423	Framework paper

The **ORCID record** (`0009-0004-0050-0264`) lists these works and stabilizes the author-entity across publications. The **Wikidata deletion discussion** for Q141392733 was also encountered, revealing that the author created a Wikidata item about himself and that its notability is contested.

The **website layer** was entered through the central DIO/ODI hub (`manifest-dio/`), the Czech and English manifesto pages, the professional biography, and the language-specific pages discoverable through site navigation.

2.2 Analytical Framework

The framework itself suggests the analytical vocabulary used throughout this report. The **DIO/ODI distinction** (applied discipline vs. theoretical meta-layer) was mapped onto the distribution ecosystem: the vocabulary and Reference Model function as the *ODI-like* semantic layer, while the manifestos, service pages, and case studies function as the *DIO-like* applied and discursive layer. The **Reconstruction Failure Modes** taxonomy—with its four analytical domains (Referential, Semantic Coherence, Temporal, Contextual)—was used as a diagnostic matrix. The **K6 semantic connectome** informed the graph perspective, though the observed distribution network is necessarily sparser and more heterogeneous than the idealized fully connected K6.

The analysis maintains a distinction between four levels of statement: (a) the observed artifact or condition—what is actually present, absent, inconsistent, accessible, or inaccessible; (b) the architectural role—how the element can be interpreted within DIO/ODI, and what reconstruction-supporting or reconstruction-straining function it has; (c) the local reconstruction consequence—what an observer entering through this representation can plausibly recover, miss, or misunderstand; and (d) the system-level significance—whether the local condition materially

compromises the identity of the framework as a whole, or whether other representations and pathways preserve sufficient continuity. The analysis does not silently jump from (a) to (d).

Where an outcome has not been directly observed, formulations such as "provides structural support for," "can be interpreted within DIO as a mechanism for," "creates a potential obstacle to," and "within this observation domain, the available representations appear sufficient to" are used. Unconditional claims such as "prevents," "ensures," or "eliminates" are avoided unless the evidence supports that scope.

The analysis distinguishes identity-preserving variation from documented conceptual development, version misalignment, incompatible current claims, susceptibility to reconstruction failure, and observed reconstruction failure. A historical edition is not treated as defective merely because it preserves an earlier state; conversely, historical provenance is not assumed to resolve every conflict encountered by a current observer.

Conceptual interdependence is kept distinct from hyperlink topology, provenance relations, repository structure, and actual observer traversal. The distribution network is not required to reproduce the K6 conceptual model literally. Numerical graph findings state their observation boundary, node and edge definitions, and relevant directionality. Acquisition failures are treated as observer-relative unless verified otherwise: a failed fetch establishes that a particular acquisition route failed, not that the resource is absent through every route.

2.3 Limitations

The **Zenodo author search interface** did not render complete results due to scheduled maintenance; the corpus was reconstructed from individual DOI records rather than from an exhaustive listing. **ResearchGate** was inaccessible. The **Wikidata discussion** was read but the item's current status is unresolved. The **graph analysis** is based on manually observed links within the collected corpus, not on a complete crawl; in-degree values represent *observed in-degree within this corpus*, not global web in-degree. The **language projections** were not compared against the Czech source, so claims about translation fidelity or semantic fragmentation remain at the level of structural possibility rather than observed condition. No **concrete observer path** through an older archived record was traced, so claims about whether an observer would encounter the reconciliation pathway remain conditional. No **downstream consumption** of the vocabulary or `llms.txt` by external systems was observed. No **independent citation or implementation analysis** was performed, so the framework's reception and reuse by external researchers or systems is not assessed.

---

3. DIO/ODI-Native Analysis

3.1 Conceptual Continuity Across Representations

The framework's core concepts exhibit **high stability at the level of definition** and **structured variation at the level of expression**. The six dimensions of the semantic connectome—Entity, Representation, Interpretation, Trust, Relationship, Reconstruction—appear consistently across the Czech manifesto, the English manifesto, the `llms.txt` `dio-odi-framework` section, the Reference Vocabulary, and the Zenodo Reference Model v0.5.0. The Reference Model explicitly describes K6 as "the complete ordered set of six mutually conditioned dimensions" and states that "K6 is not represented as a sequential pipeline or as a literal causal graph." This is a case of **genre-appropriate stabilization**: the manifesto uses metaphor (the "pentagonal pyramid," the semantic-field analogy with mass, wave, and gravitational binding) to convey the relational ontology, while the machine-readable layer encodes the same commitments as non-hierarchical, mutually related concepts. Within DIO terms, this is a form of **redundancy across media of readability**: the same conceptual structure is

encoded in natural language, in SKOS relations, in JSON-LD, and in machine-readable context files, making the core concepts recoverable from multiple entry points and genre expectations.

The **Reconstruction Failure Modes taxonomy** shows a more complex trajectory, and the Zenodo records make the version history explicit. The *Crisis* paper introduced four Entity Resolution Anomalies: Identity Collapse, Semantic Fragmentation, Temporal Discontinuity, and Semantic Vulnerability. MVDIO then introduced finer distinctions (Merge Failure, Split Failure, Temporal Misalignment, Contextual Displacement) and moved Semantic Vulnerability from a failure state to a robustness-related property. The Taxonomic Note then reorganized the full set into a **non-exclusive, multi-level hierarchy** with four analytical domains. The Reference Model v0.5.0 consolidates this: it represents the hierarchy as Referential Reconstruction Failures (Entity Substitution, Merge Failure, Split Failure, Entity Omission), Semantic Coherence Failures (Semantic Fragmentation), Temporal Reconstruction Failures (Temporal Discontinuity, Temporal Misalignment), and Contextual Adequacy Failures (Contextual Displacement), while keeping Identity Collapse as a phenomenon-level concept outside the hierarchy and Semantic Vulnerability as a robustness-related property with Semantic Robustness as its counterpart. This is not inconsistency but **versioned conceptual refinement**: later refinement does not erase the historical role of earlier publications, and each concept retains source provenance. Within DIO terms, this is a response to **Temporal Discontinuity** that preserves the historical record rather than overwriting it.

A tension emerges between the **manifesto's rhetorical ambition** ("bulletproof identity," "adversarial reading across all layers of the latent semantic space") and the **Reference Model's epistemic caution** ("conditional and non-deterministic," "without stronger causal commitment" for Reconstruction Capture). The `llms.txt` explicitly addresses this: it states that "The phrase 'bulletproof identity' expresses the strategic objective of robust identity representation across interpretation environments" and distinguishes the manifesto's positioning from the research's conceptual propositions. This is best interpreted as a **division of discursive labor**: the manifesto mobilizes and orients; the machine-readable layer constrains and specifies. The application mappings—RAG pipeline, HEM heterogeneity taxonomy, Commit-Boundary × Policy-as-Code—are explicitly marked as "derived," "non-normative," and "deliberately separated from the canonical vocabulary," preventing them from silently redefining core concepts.

Table 1: Conceptual Continuity and Divergence Across Layers

Concept / Relation	Manifesto (CZ/EN)	Reference Vocabulary / Model	`llms.txt`	Zenodo Metadata
Six dimensions (K6)	Present as fully connected graph, no arrows	Encoded as ordered set; explicitly not sequential or causal	Listed as `dio-odi-framework`	Described in v0.5.0
Reconstruction Failure Modes	Introduced as diagnostic ambition	Multi-level hierarchy, four domains; Semantic Vulnerability moved to robustness	Not enumerated in detail	Taxonomic Note and v0.5.0 encode revision
Identity Collapse	Phenomenon-level	Kept outside hierarchy	Not explicitly distinguished	Stated as phenomenon-level in v0.5.0

Concept / Relation	Manifesto (CZ/EN)	Reference Vocabulary / Model	`llms.txt`	Zenodo Metadata
Semantic Vulnerability	Risk condition	Robustness-related property, counterpart Semantic Robustness	Not explicitly mapped	Stated as robustness property in v0.5.0
Commit Boundary	Implicit in authorization ambition	Not a canonical concept; mapped to Policy-as-Code	Research topic	Referenced in related works
Bulletproof identity	Rhetorical objective	Not a canonical concept	Explicitly framed as strategic positioning	Not present in metadata

The continuity observed here is not an artifact of repetition alone. The concepts are **relationally articulated** across layers: the vocabulary's SKOS relations encode the conceptual structure that the manifesto describes in prose; the Reference Model cross-references the publications that developed the concepts; the `llms.txt` declares the evidential distinctions that the narrative pages embed implicitly. Within the inspected pathways, an observer entering through any of these layers can recover the core distinctions, though the degree of detail and the emphasis vary by genre.

3.2 Reconstruction and Its Possible Failure Modes

Applying the framework's own reconstruction vocabulary to its distribution reveals a differentiated risk profile. **Entity substitution** is structurally addressed by stable concept URIs under the canonical namespace and by persistent identifiers (ORCID, Zenodo DOIs). Within DIO terms, these URIs function as a **referential anchor**: they give observers a stable address for a concept that can be resolved across serializations and contexts. The mechanism supports reconstruction by reducing the ambiguity that would arise if concepts were identified only by natural-language labels that vary across languages and genres. Whether this anchor prevents substitution in every downstream system is not established; what is observed is that the anchor exists and is consistently referenced across the vocabulary, the Reference Model, and the web pages.

The Reference Model's preservation of the former singular "Identity-Preserving Constraint" as a separate historical record—without asserting a semantic relation to the new plural "Identity-Preserving Constraints"—is a deliberate anti-merge measure that maintains conceptual continuity while acknowledging revision. Within DIO terms, this is a **contextual-adequacy mechanism**: it sets genre expectations and prevents derived artifacts from silently redefining canonical concepts. It does not prevent an observer from conflating them, but it provides the textual basis for distinguishing them.

Semantic fragmentation is a plausible structural risk. The 16 language projections are described in the Zenodo Manifest as preserving a "shared conceptual invariant" while accommodating linguistic nuance, and the `llms.txt` states that the Czech version is the "interpretively referential" source with precedence where meanings differ. This design localizes interpretive authority but creates a structural condition in which a revision in the Czech source could, in principle, propagate unevenly across projections. The framework's own **Semantic Coherence Failures** category names this risk precisely; the architecture does not eliminate it but makes it traceable through provenance metadata and the `llms.txt` `interpretive-priority` declaration. Whether fragmentation has occurred is not observed in this analysis, because the translations were not compared against the Czech source. The risk is **plausible but unobserved**.

Temporal discontinuity is visible in the version history: the Manifest is v1.0.0 (July 2026), the Reference Model moved from v0.3.1 (October 11) to v0.5.0 (October 25), and MVDIO and the Taxonomic Note were published between these releases. The Reference Model v0.5.0 states that it "operates above the distributed corpus" and that "later refinement does not erase the historical role of earlier publications." This is a **managed temporal architecture**, not a failure, but it creates a condition in which an observer encountering an older archived record may reconstruct a different conceptual structure than one encountering the current Reference Model. The `llms.txt` `version-distinction` field addresses this directly: "Live web texts and archived releases represent their respective publication states; subsequent web revisions do not modify an already archived edition" and "Use the publication record and version of the cited archived edition." The pathway to reconciliation exists through the consolidation node, but whether an observer traverses it depends on their entry point and diligence.

Contextual displacement arises when the same concept is interpreted through different genre expectations. The `llms.txt` performs explicit genre-boundary work: it distinguishes the professional biography (first-person experience) from client references (professional relationships), case studies (practical work), and research publications (conceptual and technical propositions). This is a **pre-emptive contextual-adequacy measure** that the narrative website alone does not provide. It does not guarantee that an observer will respect the distinctions, but it provides the textual basis for doing so.

3.3 Genesis, Context, and Application

The framework's genesis is explicitly **ex-post and autoethnographic**. The Czech manifesto states that DIO "was not created as an academic derivative of existing research nor as pure abstraction" but as "an ex-post articulation of real optimization practice," methodologically grounded in "reflexive observation" with features of "autoethnography and grounded theory." The professional biography traces a trajectory from retail and journalism through psychology and communication studies to SEO and copywriting, and it describes the creation of a "multimodal E-E-A-T object" as an "experimental space" in which new DIO procedures were formed and verified. The `llms.txt` reinforces this by listing the biography as the `primary-professional-profile` and stating that it is the "Primary source for first-person experience (E-E-A-T)."

The Zenodo corpus then performs a **second-order formalization**. The Manifest dataset establishes the canonical archival edition. MVDIO performs an "adversarial reduction" that distinguishes three often-conflated levels: the existence of a reconstruction mechanism, the existence of an optimizable digital identity problem, and deliberate DIO intervention. The Taxonomic Note reorganizes the failure vocabulary. The Reference Model v0.5.0 consolidates the semantic layer. These are not merely sequential publications; they are **layered formalizations** that progressively stabilize the framework's conceptual vocabulary. The application mappings extend this into technical domains without redefining the core: the RAG mapping projects failure modes onto retrieval-augmented generation pipeline loci, the HEM mapping links heterogeneity conditions to reconstruction consequences, and the Policy-as-Code mapping translates authorization obligations into OPA/Rego and Cedar capabilities. The Reference Model's explicit statement that these mappings are "non-normative" and "deliberately separated from the canonical vocabulary" maintains the boundary between descriptive ontology and normative application.

Within DIO terms, this genesis and formalization trajectory can be interpreted as a **transition from lived professional experience to stable semantic artifacts**. The framework's authority claim is not primarily academic-theoretical but practitioner-reflexive; the vocabulary and Reference Model then perform a second-order formalization that converts that experience into concepts with stable identifiers and relations. This trajectory is itself a form of **reconstruction**: the framework reconstructs its own genesis as a coherent development from practice to theory to machine-readable specification. Whether an observer entering through the vocabulary alone can recover this genesis is

not established; the vocabulary provides the concepts, but the narrative of their emergence is recoverable primarily through the manifesto and biography.

3.4 Multilingual Representation

The language inventory is explicitly documented: the Czech original plus 15 additional projections (EN, DE, FR, ES, IT, PT, PL, FI, HU, UK, ZH, JA, KO, AR, HI), yielding **16 language codes**. The Zenodo Manifest record describes the design intent: the projections "strictly preserve a shared conceptual invariant" while accommodating "specific cultural and linguistic nuances." The `llms.txt` states that the Czech version is the `canonical-language` and that "the Czech canonical source takes interpretive precedence where language versions differ in meaning." The hub (`manifest-dio/`) is described as the "primary nodal point of language projections."

This is a **centripetal and asymmetric** topology. Within DIO terms, the Czech source functions as a **nodal point**: a privileged signifier that partially fixes meaning across the field of linguistic projections. The architecture supports a shared conceptual core but creates distinct interpretive contexts: a reader of the Arabic or Japanese projection encounters the framework through a translation whose conceptual neighbors may not be symmetrically cross-linked. The `llms.txt` `manifesto` section acknowledges this: "Additional language projections are linked from the manifesto hub" and "interpretive priority: The Czech canonical source takes interpretive precedence." The framework's own **Semantic Fragmentation** concept illuminates the residual risk: fragmentation is not an absence of translation but a structural possibility that arises when the translation network is hub-dependent and version asynchrony is possible. The `llms.txt` `version-handling` instruction—"Preserve distinctions between live pages, archived publications, language versions, and reference artifacts"—is a partial mitigation, but it relies on the observer's compliance. Whether fragmentation has actually occurred is not observed, because the translations were not compared against the Czech source.

3.5 Machine-Readable Representation and Non-Human Observers

The machine-readable layer performs functions that the narrative layer cannot. The **SKOS vocabulary** makes explicit the broader/narrower and related-concept relations that the manifesto expresses through metaphor and prose. The **JSON-LD and Turtle serializations** enable retrieval systems, knowledge graphs, and ontology tools to identify DIO/ODI concepts by stable URI and traverse their semantic relations without natural-language interpretation. The **GitHub repository** provides versioned release artifacts and documentation explaining design choices—most notably the SKOS-first strategy and the deferral of OWL axioms.

The `llms.txt` file adds a declared interface for non-human observers. Its sections (`identity`, `eeat`, `key-pages`, `dio-odi-framework`, `manifesto`, `research`, `reference-artifacts`, `ai-usage-policy`, `crawling-priorities`) make explicit what the narrative pages leave implicit: the author's identity fields, the evidential distinctions between professional and research claims, the versioned nature of the corpus, and the permissions and restrictions for AI use. The `llms-full.txt` goes further by providing a snapshot of eight selected pages with explicit transformation metadata (diacritics removed, ASCII-only, reversible Unicode notation for non-Latin scripts). This is a **self-documenting export**: it declares its own scope, its source documents, and its limitations ("not the entire website or research corpus").

In the framework's own terms, `llms.txt` functions as a **declared interface** for non-human observers: it provides prioritized, structured context that can support machine interpretation and retrieval. The file is designed to make explicit the evidential distinctions and version boundaries that the narrative pages leave implicit. It is interpretively convenient—it compresses the site's conceptual structure and assigns explicit priority levels to pages—but it is **not structurally necessary for reconstruction**. The equivalent distinctions remain recoverable from the vocabulary (which distinguishes canonical concepts from application mappings), the publication records (which distinguish versions and

provenance), the repository documentation (which explains design choices), and the narrative corpus (the biography states its first-person character; the manifesto states its version status; the Reference Model states its non-normative relationship to application mappings). If `llms.txt` were unavailable, an observer would lose a compressed, prioritized interface, but not the underlying conceptual structure.

The `llms.txt` `ai-usage-policy` declares `training: allow`, `citations: require`, `persona: deny`, and `style-imitation: deny`. These are **stated preferences and permissions**, not technically enforced access-control constraints. The file's authority derives from the author's declaration, not from a verifiable authorization protocol. This is consistent with the Reference Model's caution that it "does not replace the papers" and "does not claim to be a completed formal ontology." Within the framework's own terms, the `llms.txt` can *constrain interpretation* by providing stable definitions and relations, but it cannot *authorize* a reconstruction outcome in the sense developed in the Commit Boundary paper. The authorization analogy reaches its limit precisely at the point where machine-readable artifacts shift from being descriptive aids to normative gatekeepers. The `llms.txt` policy's effectiveness depends on the voluntary compliance of AI systems; its presence does not ensure compliance.

A more significant question is whether the file's **presence** creates a new dependency: if an LLM relies on `llms.txt` and the file diverges from the live pages, the file could become a **contextual displacement vector** rather than a mitigation. The `llms-full.txt` export metadata mitigates this by stating its snapshot date and scope, but the live `llms.txt` does not carry an explicit version marker in the inspected content. This is a **potential** strain, not an observed one.

---

4. Graph Report

4.1 Chosen Graph Model and Observation Boundary

The ecosystem was modeled as a **heterogeneous directed network** with distinct node types (web pages, vocabulary concepts, repository artifacts, archival records, identity profiles, context files) and edge types (hyperlinks, citation and provenance relations, semantic relations, language-alternate declarations, authorship assertions). Separate views are maintained for the machine-readable layer, the archival layer, and the web-discursive layer, while cross-layer bridges are examined. The **observed boundary** consists of the resources directly inspected or returned in search results during this research. In-degree values are *observed in-degree within the collected corpus*, not global web in-degree.

The conceptual and infrastructural graphs are kept distinct. The K6 conceptual model is not required to appear literally in the hyperlink topology. Conceptual interdependence—the mutual conditioning of the six dimensions—is a theoretical claim about digital identity; the hyperlink and provenance network is an empirical structure observed in the distribution ecosystem. The analysis does not conflate the two.

4.2 Two Parallel Entry Points: Zenodo vs. Web Hub

The comparison uses two entry points that were actually inspected: the Zenodo archival record (via DOI resolution and ORCID) and the website hub (`manifest-dio/`). The comparison is limited to the inspected pathways; it does not claim to exhaust the discoverable resources from each entry point.

Table 2: Entry-Point Comparison

Dimension	Zenodo Entry Point	Web Hub Entry Point
Starting resource	ORCID record → DOI records → Manifest, Reference Model, MVDIO, Taxonomic Note, Crisis, Conceptual Framework	`manifest-dio/` → Czech/English manifestos, language projections, bio, service pages, `llms.txt`
What is reached	Publication-centric reconstruction: versioned datasets, conceptual papers, reference model, provenance relations	Concept-centric reconstruction: manifesto, language projections, genesis narrative, professional evidence, machine-readable context
What remains peripheral	Individual language projections; service/application pages; `llms.txt`	Archival version history; DOI relations; Reference Model releases
Explicit relations	`IsDerivedFrom`, `IsIdenticalTo`, `References`; ORCID work list	Language alternates; links to bio, vocab, service pages; `llms.txt` crawling priorities
Reconstruction strengths	Strong provenance and version awareness; explicit lineage from Manifest to later works; stable author-entity via ORCID	Strong conceptual orientation; multilingual reach; genesis narrative; explicit evidential distinctions via `llms.txt`
Reconstruction limitations	Does not expose the multilingual distribution or the professional evidence base	Does not systematically expose archival version history or DOI relationships
Cross-layer bridge	Links to vocabulary endpoint; Reference Model v0.5.0 consolidates the corpus	Links to vocabulary endpoint; `llms.txt` lists reference artifacts

The two entry points produce **complementary but non-identical reconstructions**. Neither alone provides the full picture. The Zenodo entry point is stronger for **temporal and referential reconstruction**: it gives an observer the dated publication sequence, the provenance relations, and the version history. The web hub is stronger for **conceptual and contextual reconstruction**: it gives an observer the framework's self-description, its multilingual reach, and its genesis narrative. The **Reference Model v0.5.0** functions as a cross-layer bridge: it is a Zenodo record that explicitly describes its relationship to the distributed corpus, and it is linked from the vocabulary. An observer who enters through either point and follows the available links can reach the other. The pathways exist, but they are not symmetrical: the web hub does not systematically link to the Zenodo records, while the Zenodo records do link to the vocabulary and to external preprints.

Within DIO terms, this is a **partially meshed network with a cross-layer consolidation node**. The consolidation node supports reconstruction by providing a point from which both the archival and the conceptual layers can be reached. The residual strain is that an observer who does not traverse the consolidation node may reconstruct a version of the framework that is either publication-centric without the conceptual overview, or concept-centric without the version history.

4.3 Node Degree, Centrality, and Topology

Table 3: Observed Structural Positions (Manual Observation, Not Exhaustive Crawl)

Node	Observed In-Degree (within corpus)	Observed Out-Degree (within corpus)	Structural Role
`manifest-dio/` (web hub)	High (linked from bio, manifesto, language pages, `llms.txt`)	High (links to CZ/EN manifestos, bio, language projections)	Central navigational hub
Zenodo Manifest dataset (21610934)	High (referenced by MVDIO, Crisis, v0.5.0, `llms.txt` citation helper)	Moderate (describes distribution layers)	**Archival root node**
`vocab.danielberanek.cz/dio-ontology/`	Moderate–High (linked from GitHub, Reference Model, web pages)	Moderate (links to mappings, repository)	**Primary semantic bridge**
Zenodo Reference Model v0.5.0 (23643845)	Low–Moderate (newest record)	High (describes vocabulary, mappings, related works)	**Consolidation node**
GitHub repository	Moderate (linked from vocabulary, Reference Model)	Moderate (links to releases, concept map)	Provenance and release bridge
ORCID record	Low–Moderate (linked from Zenodo author fields)	High (lists all works)	Author-entity anchor
`llms.txt`	Low (new artifact; not yet widely linked)	High (describes site structure, policies)	**Machine-readable interface**
Czech manifesto	High (linked from hub and `llms.txt`)	Moderate (links to hub, bio)	Interpretive reference node
English manifesto	High (linked from hub and `llms.txt`)	Moderate (links to hub)	Principal non-Czech entry
Individual language projections	Low (linked mainly from hub)	Low (link back to hub)	Peripheral spokes
Application mapping pages	Low–Moderate (linked from vocabulary and v0.5.0)	Low (link to vocabulary)	Specialized extensions

The observed topology is **hub-and-spoke with partial meshing**. The web hub is the most connected node in the web layer, but the **Zenodo Manifest dataset** is the most important **archival root**: MVDIO and the Crisis paper both declare `IsDerivedFrom` the Manifest, and the Reference Model v0.5.0 cross-references the Manifest and the subsequent works. The vocabulary remains the most important **cross-layer semantic bridge**, linked from the GitHub repository, the Reference Model, and the web pages. The `llms.txt` is a **machine-readable interface node** whose centrality is difficult to measure because it is a recently added artifact, but its structural role is high: it declares the entire site's architecture, evidential distinctions, and AI policies in a single file.

The multilingual subgraph is **centripetal**: the Czech original and the hub have the highest in-degree, while individual language projections are peripheral spokes. This is consistent with the stated design ("Czech canonical source takes interpretive precedence"), but it means the multilingual network does not provide **redundant paths** to core concepts. If the hub or the Czech source is unavailable, the language projections become isolated interpretive contexts.

4.4 Cross-Layer Connections and Bottlenecks

The strongest cross-layer connections are **identifier-based**: the ORCID iD connects the author-entity across publications, the Zenodo concept DOI connects the Manifest across versions, and the canonical namespace connects concepts across serializations. The weakest cross-layer connections are **hyperlink-based**: the Zenodo records do not consistently link to individual language projections, and the web hub does not consistently link to DOI-versioned Reference Model releases. The `llms.txt` partially addresses this by listing `https://vocab.danielberanek.cz/dio-odi/` and the GitHub repository under `reference-artifacts`, but it does not enumerate the Zenodo records or their DOIs.

The most significant **structural vulnerability** is the **single-operator dependency**. The ORCID record, the website's biography, the `llms.txt` `identity` section, and the Zenodo author fields all converge on Daniel Beránek as the sole author and maintainer. Within the observed snapshot, the artifacts are maintained and mutually consistent; the dependency is a **future-oriented risk**, not an observed reconstruction failure. If the author ceased maintenance, the live vocabulary, the hub, and the `llms.txt` would gradually diverge from the archived record. The archived Zenodo records would remain accessible, but the cross-layer bridges (the live vocabulary, the hub) might not. The framework's own **Reconstruction Capture** concept—a related concept without stronger causal commitment in the Reference Model—is relevant here as a *hypothesis*: the centrality of a single author-operator could, under certain conditions, capture the reconstruction process around a single perspective, reducing the diversity of interpretive paths. This is not an observed effect; it is a structural condition that the framework's own vocabulary can name.

A second bottleneck is the **Wikidata deletion discussion**. The author created a Wikidata item (Q141392733) about himself, explicitly disclosing a conflict of interest and arguing that the item serves a "bibliographic structural function" as an author node for the DOI-linked works. The discussion is unresolved. This indicates that reconstruction within the open knowledge-graph infrastructure is not fully secured. The author-entity remains stabilized by ORCID and Zenodo; the Wikidata question is material to the framework's **reach** into the open knowledge-graph infrastructure, but not to its **core reconstructability**. Within DIO terms, it is a **Contextual Displacement** risk: the entity may be robustly reconstructable within the author's own distribution network but not within Wikidata's notability regime.

---

5. Materiality Assessment of Principal Weaknesses

The following assessment evaluates the materiality of the principal weaknesses for the reconstruction task. The standard is not whether a defect exists, but whether it obstructs the inspected reconstruction pathways. For each weakness, the assessment identifies the reconstruction task affected, the observer or entry point affected, whether the difficulty is locally contained or propagates across layers, and whether another accessible representation contextualizes, corrects, or compensates for it. The assessment does not assume that redundancy guarantees compensation, nor that local failure defeats the distributed whole.

5.1 Multilingual Hub-Dependence

Observed condition. The language inventory comprises 16 language codes. The Czech version is declared the interpretive reference. Individual language projections are linked from the central hub but do not appear to be symmetrically cross-linked with one another.

Architectural role within DIO. This is a **nodal-point architecture**: the Czech source functions as a privileged signifier that partially fixes meaning across the field of projections. It is consistent with the framework's own theory of nodal points as points of partial fixation within a discursive field.

Local reconstruction consequence. An observer entering through a non-Czech projection can recover the core conceptual structure (the six dimensions, the DIO/ODI distinction, the failure-mode taxonomy) because those are restated in the projection. The observer may not recover the full provenance of a specific translation choice, and may not know which version of the Czech source the projection corresponds to, unless they navigate back to the hub. The pathway back to the hub exists.

System-level significance. This strain is **locally contained**. The framework's identity does not depend on full multilingual symmetry; it depends on the core concepts being recoverable, which they are. The risk of semantic fragmentation across translations is **plausible but not observed**: the translations were not compared against the Czech source, so whether fragmentation has occurred cannot be stated. What can be stated is that the architecture creates the structural condition for fragmentation if a revision propagates unevenly, and that the `llms.txt` version-handling instruction is designed to mitigate this by directing observers to preserve distinctions between versions.

5.2 Version Misalignment Across Archived Releases

Observed condition. The vocabulary is at v0.2.0; the Reference Model moved from v0.3.1 to v0.5.0; earlier archived versions preserve earlier vocabulary states (v0.4.0 preserves v0.1.0). The Reference Model v0.5.0 explicitly documents the transition from the singular to the plural "Identity-Preserving Constraints" and preserves the singular as a historical record.

Architectural role within DIO. This is a **managed temporal architecture**: the archival layer preserves earlier states, and the current consolidation node documents the transition. Within DIO terms, it is a response to **Temporal Discontinuity** that does not erase the historical record.

Local reconstruction consequence. An observer entering through an older archived record may encounter a conceptual state that differs from the current Reference Model. For example, an observer encountering v0.1.0 vocabulary may not know that v0.2.0 introduced a separate plural concept. The pathway to reconciliation exists through the v0.5.0 Reference Model, which explicitly documents the transition. Whether an observer actually traverses that pathway depends on their entry point and their diligence.

System-level significance. This strain is **locally contained but not eliminated**. The framework's identity is not compromised by the existence of earlier versions; it is compromised only if an observer cannot distinguish a historical state from a current one. The Reference Model's explicit documentation of the transition provides the basis for that distinction. The residual risk is that an observer entering through an older record may not know to look for the consolidation node. This is a **plausible but unobserved** reconstruction failure: no concrete observer path through an older record was traced. The materiality is therefore **conditional**: if the pathway is traversed, the strain is contained; if it is not, the observer may reconstruct an earlier state as if it were current.

5.3 Single-Operator Dependency

Observed condition. The ORCID record, the website biography, the `llms.txt` identity section, and the Zenodo author fields all converge on Daniel Beránek as the sole author and maintainer. The GitHub repository is under a single account. The vocabulary is hosted on a domain controlled by the author.

Architectural role within DIO. This is a **centralized maintenance structure**. Within DIO terms, it concentrates the **Trust** and **Relationship** dimensions in a single entity: the author is the primary source of provenance, the primary maintainer of the vocabulary, and the primary point of contact for the framework.

Local reconstruction consequence. An observer can reconstruct the framework from the current artifacts without needing to interact with the author. The dependency is on the author's **continued**

maintenance, not on the author's **presence** for any particular reconstruction act. The current artifacts are internally consistent and mutually referential.

System-level significance. This is a **structural vulnerability, not an observed failure**. The ecosystem currently holds together because the artifacts are maintained and consistent. The vulnerability is that if the author ceased maintenance, the live vocabulary, the hub, and the `llms.txt` would gradually diverge from the archived record. The archived Zenodo records would remain accessible, but the cross-layer bridges (the live vocabulary, the hub) might not. Within the observed snapshot, this vulnerability does not materially compromise reconstruction because the bridges are currently intact. It is a **future-oriented risk**, not a present reconstruction failure.

5.4 Wikidata Notability Question

Observed condition. A Wikidata item (Q141392733) about the author exists, and its notability is contested in a deletion discussion. The author disclosed a conflict of interest and argued that the item serves a bibliographic structural function as an author node for DOI-linked works.

Architectural role within DIO. Wikidata is a **knowledge-graph infrastructure** that the framework analyzes as one of the media of readability. The author item, if retained, would provide a stable node in that infrastructure linking the author-entity to the publications. If deleted, the author-entity remains stabilized by ORCID and Zenodo, but not by Wikidata.

Local reconstruction consequence. An observer entering through Wikidata would either find the item (and from it, potentially, the publications) or not. If the item is deleted, the observer would need to enter through another route (ORCID, Zenodo, the website). The pathway through ORCID and Zenodo exists.

System-level significance. This is a **contextual-adequacy strain, not a referential failure**. The framework's identity does not depend on Wikidata retention; it depends on the author-entity being reconstructable, which it is through ORCID and Zenodo. The Wikidata question is material to the framework's **reach** into the open knowledge-graph infrastructure, but not to its **core reconstructability**. Within DIO terms, it is a **Contextual Displacement** risk: the entity may be robustly reconstructable within the author's own distribution network but not within Wikidata's notability regime.

5.5 `llms.txt` Unavailability

Observed condition. The `llms.txt` file provides a structured, machine-readable summary of the site's identity, evidential distinctions, crawling priorities, and AI usage policy. It is a relatively recent addition and is not referenced from the Zenodo records or the vocabulary.

Architectural role within DIO. It is a **declared interface for non-human observers**: it compresses the site's conceptual structure and makes explicit distinctions that the narrative pages leave implicit. It is designed to support machine interpretation and retrieval.

Local reconstruction consequence. If `llms.txt` were unavailable, an observer (human or machine) would lose a convenient, prioritized summary. The core conceptual distinctions—DIO vs. ODI, canonical vs. derived, live vs. archived, professional vs. research evidence—are also recoverable from the narrative corpus, the Reference Model, and the publication metadata. The biography states its first-person, experiential character; the manifesto states its version status; the Reference Model states its non-normative relationship to application mappings.

System-level significance. This is **not a material weakness**. The `llms.txt` file is interpretively convenient but not structurally necessary. The equivalent distinctions remain recoverable from the vocabulary, the publication records, the repository documentation, and the narrative corpus. The file's absence would reduce the **efficiency** of machine-mediated reconstruction, not its **possibility**.

The more significant question is whether the file's **presence** creates a new dependency: if an LLM relies on `llms.txt` and the file diverges from the live pages, the file could become a **contextual displacement vector**. The `llms-full.txt` export metadata mitigates this by stating its snapshot date and scope, but the live `llms.txt` does not carry an explicit version marker in the inspected content. This is a **potential** strain, not an observed one.

Table 4: Materiality Assessment Summary

Weakness	Reconstruction Task Affected	Observer/Entry Point Affected	Locally Contained?	Compensation Inspectable?	System-Level Significance
Multilingual hub-dependence	Semantic coherence across projections	Non-Czech language entry points	Yes	Core concepts restated in projections; hub pathway exists	Locally contained; plausible but unobserved fragmentation risk
Version misalignment	Temporal reconstruction	Older archived record entry points	Yes, conditionally	Consolidation node documents transitions; pathway depends on traversal	Locally contained but not eliminated; conditional on observer path
Single-operator dependency	Continued maintenance of live artifacts	All entry points over time	Future-oriented	Archived records remain accessible; live bridges at risk	Structural vulnerability; not an observed failure at snapshot
Wikidata notability	Reach into open knowledge-graph infrastructure	Wikidata entry point	Yes	ORCID and Zenodo stabilize author-entity	Contextual-adequacy strain; not referential failure
`llms.txt` unavailability	Efficiency of machine-mediated reconstruction	AI-mediated entry points	Yes	Equivalent distinctions recoverable from other artifacts	Not material; convenience, not necessity

---

6. Emergent Insights

6.1 The Framework's Reflexive Productivity

The most consequential emergent finding is that **DIO/ODI's conceptual vocabulary is genuinely productive when turned upon its own distribution**. The Reconstruction Failure Modes taxonomy does not merely describe external entity-resolution problems; it generates a differentiated diagnosis of the framework's own representational ecosystem. The observed mitigation of entity substitution and merge failure through stable URIs and explicit conceptual separations is a **positive alignment** between theoretical commitment and implementation. The residual risks—semantic fragmentation

across languages, temporal discontinuity across versions, contextual displacement across genres—are precisely the failure modes that the framework names and partially addresses through provenance metadata, versioned releases, `llms.txt`'s evidential distinctions, and the Reference Model's explicit cross-referencing. This is not a circular vindication: the framework's own tools reveal *where* its implementation is strong and *where* it remains structurally vulnerable.

6.2 The SKOS-First Strategy as an Epistemic Position

The decision to begin with SKOS rather than OWL is not merely a technical convenience. It is a **theoretical statement about the nature of digital identity**. OWL's stronger axioms would commit the framework to a single, logically closed model of identity relations. SKOS supports broader/narrower and related-concept relations without enforcing disjointness or existential restrictions. This aligns with the framework's core claim that digital identity is **distributed, emergent, and dynamically reconstructed** rather than a fixed logical object. The Reference Model's statement that "later refinement does not erase the historical role of earlier publications" and its preservation of the former singular concept as a separate historical record are concrete expressions of this epistemic position: the ontology is designed to accommodate revision without losing provenance. This is a rare and deliberate form of **ontological humility** in a domain that often rushes toward formalization.

6.3 `llms.txt` as a Declared Interface, Not a Structural Necessity

The `llms.txt` file introduces a structural element that the web-only view does not capture. It is a **declared interface** for non-human observers: it states what the site is, what it contains, how it should be navigated, what evidence categories exist, and what AI systems are permitted to do. Its `crawling-priorities` section assigns explicit priority levels to pages, effectively **curating the reconstruction path** for an LLM. Its `ai-usage-policy` declares permissions and restrictions. Its `summary` section compresses the entire site's conceptual structure into a few paragraphs.

This is a significant development in the framework's distribution strategy: it moves beyond merely *publishing* machine-readable artifacts (vocabulary, JSON-LD) to *declaring* a machine-readable interface. But the file is **interpretively convenient, not structurally necessary**. The evidential distinctions it makes explicit are also recoverable from the vocabulary, the publication records, the repository documentation, and the narrative corpus. If the file were unavailable, an observer would lose a compressed, prioritized interface, but not the underlying conceptual structure. The more significant question is whether the file's **presence** creates a new dependency: if an LLM relies on `llms.txt` and the file diverges from the live pages, the file could become a **contextual displacement vector** rather than a mitigation. The `llms-full.txt` export metadata is careful to state that it is a snapshot of eight selected pages, not the entire corpus, and that "live pages and separately archived publications have their own version histories." This self-limitation is a good-faith effort to manage the temporal discontinuity risk, but it also underscores that the `llms.txt` layer is **not a canonical source**; it is a derived, author-maintained summary.

6.4 The Wikidata Tension

The Wikidata deletion discussion illuminates the framework's **identity infrastructure dependencies**. The author's argument for retaining the Wikidata item is explicitly structural: the item serves as an author node linking a set of DOI-identified works. The deletion discussion raises notability concerns. This is a **real-world test of the framework's reconstruction thesis**: can an entity be coherently reconstructed across the open knowledge-graph infrastructure when its notability is contested? The framework's own **Contextual Adequacy Failures** category—specifically **Contextual Displacement**—is the relevant diagnostic. The entity may be robustly reconstructable within the author's own distribution network (website, Zenodo, ORCID, vocabulary) but not within Wikidata's notability regime. This is not a failure of the framework's concepts; it is a demonstration that **reconstruction**

depends on the norms of the observing infrastructure, and those norms are not under the author's control.

6.5 Machine-Readable Artifacts and the Limits of Authorization

The `llms.txt` `ai-usage-policy` declares `training: allow`, `citations: require`, `persona: deny`, and `style-imitation: deny`. These are **stated preferences**, not technically enforced constraints. The vocabulary and Reference Model can *constrain interpretation* by providing stable definitions and relations, but they cannot *authorize* a reconstruction outcome. The Commit Boundary paper's central distinction—between evidence available for reasoning and evidence eligible to authorize a persistent identity-changing mutation—applies here only by analogy. A machine-readable vocabulary can provide *evidence* for reconstruction; it cannot *authorize* a commit. The `llms.txt` policy's authority derives from the author's declaration, and its effectiveness depends on the voluntary compliance of AI systems. This is a **category boundary** that the framework correctly maintains: the distribution artifacts are descriptive and constraining, not normative and enforcing. The authorization distinction is subordinate to the reconstruction question: what matters for reconstruction is whether the artifacts provide sufficient evidence for an observer to recover the framework's identity, not whether they can enforce a particular interpretive outcome.

---

7. Synthetic Conclusion

When DIO/ODI's own conceptual apparatus is applied to its implementation and distribution at the observed snapshot, the framework emerges as **internally coherent, structurally layered, and reflexively productive**. The strongest alignments are between the framework's theoretical commitment to distributed, reconstructable identity and its implementation through stable concept URIs, a versioned Zenodo corpus with explicit provenance relations, and a SKOS-first vocabulary that refuses premature logical closure. The machine-readable layer does not replace the narrative corpus; it provides a semantic anchor that allows the corpus to be navigated as an evolving, versioned conceptual object. The `llms.txt` adds a declared interface for non-human observers, making explicit the evidential distinctions and version boundaries that the narrative pages leave implicit—though these distinctions are also recoverable from other artifacts.

What appears to hold together. The core conceptual structure—the DIO/ODI distinction, the six dimensions of the semantic connectome, and the Reconstruction Failure Modes taxonomy—is recognizably continuous across the vocabulary, the Reference Model, the manifesto, the `llms.txt`, and the application mappings. The stable concept URIs provide a referential anchor. The Zenodo provenance relations provide a temporal anchor. The Reference Model v0.5.0 provides a consolidation node that cross-references the distributed corpus. The Czech manifesto provides an interpretive reference for the language projections. Within the inspected pathways, an observer entering through either the Zenodo record or the web hub can recover the framework's core identity, even if the two entry points emphasize different aspects.

Where continuity is strained. The multilingual network is centripetal: the Czech source and the hub carry the interpretive authority, and the individual language projections are peripheral spokes. The version history creates conditions in which an observer entering through an older record may encounter a different conceptual state; the Reference Model documents the transitions, but the pathway to reconciliation depends on the observer traversing it. The single-operator dependency is a structural vulnerability: the current artifacts hold together, but their continued coherence depends on maintenance that is concentrated in one person. The Wikidata notability question indicates that the author-entity is not fully stabilized within the open knowledge-graph infrastructure, although it is stabilized within the scholarly infrastructure (ORCID, Zenodo). The `llms.txt` is interpretively

convenient but not structurally necessary; its absence would reduce the efficiency of machine-mediated reconstruction, not its possibility.

What cannot be determined from the observed snapshot. Whether the language projections are semantically faithful to the Czech source (the translations were not compared). Whether an observer entering through an older archived record would actually encounter the reconciliation pathway (no concrete observer path was traced). Whether the `llms.txt` file is consumed by LLMs as intended (no downstream consumption was observed). Whether the Wikidata item will be retained (the discussion is unresolved). Whether external systems actually use the vocabulary or the application mappings (no independent citation or implementation analysis was performed). These are open questions that would require additional investigation—translation comparison, observer-path tracing, usage-log analysis, and citation analysis—rather than further interpretation of the observed artifacts.

Overall assessment. Within the observed snapshot and the inspected pathways, the DIO/ODI ecosystem appears to sustain a recognizable and meaningfully reconstructable identity. The weaknesses are real but locally contained: they create conditions under which reconstruction *could* fail for a particular observer entering through a particular route, but the ecosystem provides multiple pathways to the core conceptual structure, and the most consequential distinctions are restated across independent artifact types. The framework's own concepts—referential anchoring, provenance relations, nodal points, contextual adequacy—provide a differentiated account of why this is so, and of where the residual strains remain. The assessment is neither an unconditional endorsement nor a mechanical compliance audit. It is a reflexive analysis conducted through DIO/ODI's own conceptual perspective, and its conclusion is that the distributed identity holds together at the observed snapshot, with strains that are locally contained and vulnerabilities that are future-oriented rather than presently realized.

Source Citations

- `llms.txt` and `llms-full.txt`, supplied by the user (snapshot date 2026-10-29). Key sections: `identity`, `eeat`, `manifesto`, `research`, `reference-artifacts`, `ai-usage-policy`, `crawling-priorities`, `export-metadata`.
- Zenodo Manifest dataset, DOI 10.5281/zenodo.21610934, v1.0.0, 2026-07-26.
- Zenodo Reference Model v0.5.0, DOI 10.5281/zenodo.23643845, 2026-10-25.
- Zenodo Reference Model v0.3.1, DOI 10.5281/zenodo.23290683, 2026-10-11.
- Zenodo MVDIO, DOI 10.5281/zenodo.23226205, v1.0.0, 2026-10-08.
- Zenodo Taxonomic Note, DOI 10.5281/zenodo.23269323, 2026-10-09.
- Zenodo *Crisis of Entity Resolution*, DOI 10.5281/zenodo.22257504, v1.0.0, 2026-09-02.
- ORCID record, `https://orcid.org/0009-0004-0050-0264`.
- Reference Vocabulary, `https://vocab.danielberanek.cz/dio-odi/`.
- GitHub repository, `https://github.com/dan-beranek/dio-odi-reference-model`.
- Wikidata deletion discussion for Q141392733.

Authorship and AI use

This analysis was conducted within a research project conceived and directed by Daniel Beránek. The author defined the research question, selected the initial research entry points, and specified the analytical perspective and research requirements. DeepSeek carried out the analysis and generated the report in response to the author's integrated assignment. The documented assignment is reproduced in full in the final appendix.

The designation “DeepSeek” identifies the system used; the precise version of its underlying language model is not specified.

Appendix: Author-Designed Research and Revision Prompts

Initial research prompt

deep-research ROLE & OBJECTIVE You are an interdisciplinary researcher in Knowledge Representation, Semantic Web, Digital Identity, Graph Theory, and strategic communication. Conduct a multidimensional deep research analysis of the implementation, distribution, and semantic coherence of Digital Identity Optimization (DIO) and Ontology of Digital Identity (ODI), authored by Daniel Beránek. The central research question is: What does the implementation and distribution of DIO/ODI reveal when examined through the framework’s own conceptual logic, complemented by graph-theoretical analysis? Explore how the framework is represented across machine-readable artifacts, archival publications, multilingual web texts, and identity-linked platforms. Investigate how these representations relate, where they reinforce or complicate one another, and what their structure means for the reconstruction of the framework and its author. INITIAL ORIENTATION: MACHINE-READABLE LAYER FIRST Begin strictly with these two resources: Reference Vocabulary: https://vocab.danielberanek.cz/dio-odi/ Corresponding repository: https://github.com/dan-beranek/dio-odi-reference-model Before entering the broader narrative corpus, inspect the available machine-readable vocabulary, concept relations, schema, repository structure, and relevant documentation. Follow links within these resources to locate the underlying machine-readable artifacts rather than relying only on their introductory pages. The purpose of this sequence is to establish a conceptual orientation before encountering the volume and rhetorical density of the surrounding texts. Reconstruct the framework’s terminology and internal relationships from this layer first. Once this initial orientation is established, proceed to the archival and web representations. Revisit your initial understanding as the wider corpus reveals historical development, additional context, or tensions. RESEARCH FREEDOM Use your understanding of DIO/ODI to develop your own research approach. Let the framework suggest relevant questions, analytical perspectives, connections, and methods beyond those explicitly listed below. Where useful, draw on graph theory, discourse analysis, cybernetics, information architecture, or other approaches that emerge from the material. Treat the proposed tasks as starting points rather than an exhaustive checklist. You may refine their assumptions, follow unexpected findings, and identify relationships that the assignment does not anticipate. Evaluate DIO/ODI through its own conceptual logic without presupposing that its implementation either succeeds or fails. Keep clear what you observed, what you inferred, and what remains unresolved. STARTING CORPUS & FUNCTIONAL LAYERS LAYER 1: MACHINE-READABLE REFERENCE LAYER Reference Vocabulary: https://vocab.danielberanek.cz/dio-odi/ Source repository: https://github.com/dan-beranek/dio-odi-reference-model Explore the vocabulary, persistent concept identifiers, semantic relations, available serializations, application mappings, and version history. Examine how the published vocabulary and repository relate to one another and to the rest of the ecosystem. LAYER 2: ARCHIVAL, PUBLICATION & IDENTITY INFRASTRUCTURE Zenodo author corpus: https://zenodo.org/search?q=metadata.creators.person_or_org.name%3A%22Ber%C3%A1nek%2C%20Daniel%22&l=list&
ORCID: https://orcid.org/0009-0004-0050-0264 ResearchGate: https://www.researchgate.net/profile/Daniel-Beranek Explore the relevant publications, metadata, identifiers, versions, attribution links, and relationships between records. Investigate the distinct roles these resources play in persistence, discoverability, authorship, provenance, distribution, and trust. LAYER 3: WEB & DISCURSIVE REPRESENTATIONS Root domain: https://danielberanek.cz Central DIO/ODI hub: https://danielberanek.cz/manifest-dio/ Professional biography and genesis-related source: https://danielberanek.cz/od-pultu-az-k-seo-a-copy-s-ostrymi-hroty-profesni-bio/ Czech

manifesto: https://danielberanek.cz/manifest-digital-identity-optimization-dio/ English manifesto: https://danielberanek.cz/manifesto-of-digital-identity-optimization-dio-and-ontology-of-digital-identity-odi/ AI-oriented context files: https://danielberanek.cz/llms.txt https://danielberanek.cz/llms-full.txt Discover the other language versions and relevant pages through the website's links and metadata. Consider the website as a whole where its broader structure contributes to the representation of DIO/ODI or its author. These URLs are starting points, not a closed corpus. TASK 1: DIO/ODI-NATIVE METAMETHODOLOGICAL EXPLORATION Apply the conceptual and analytical resources of DIO/ODI to its own implementation and distribution. Develop the evaluation from your understanding of the framework rather than merely applying an external checklist. Where the framework suggests an analytical method without explicitly specifying one, explain how you derived it. Explore the following questions while remaining open to others. CONCEPTUAL CONTINUITY ACROSS REPRESENTATIONS How are the framework's core concepts and relationships expressed in: the reference vocabulary; archival publications; web texts; language versions; application mappings; AI-oriented context files? Investigate what remains stable, what changes, and how the representations complement or challenge one another. Consider whether differences reflect genre, audience, translation, historical development, conceptual revision, or substantive inconsistency. RECONSTRUCTION & ITS POSSIBLE FAILURE MODES Use the reconstruction-related concepts you find in DIO/ODI to examine the ecosystem. Potential questions include: How are the author, professional persona, framework, publications, and applications distinguished and connected? Where might entity substitution, merging, splitting, omission, or attribution ambiguity arise? Does distribution across platforms and languages support a coherent reconstruction, produce fragmentation, or do both under different conditions? Which contextual transitions preserve meaning, and which create interpretive tension? Distinguish observed reconstruction problems from structural conditions that might enable them. GENESIS, CONTEXT & APPLICATION Explore the relationship between the framework's genesis, the author's professional biography, and its later theoretical and technical developments. How does the transition between lived professional experience, strategic communication, digital identity theory, and technical applications affect the framework's interpretation? Examine relevant application mappings and papers discovered in the corpus. Ask whether they clarify the core model, extend it, reinterpret it, or expose unresolved questions. MULTILINGUAL REPRESENTATION Discover the actual language inventory and examine how its representations relate. Investigate: conceptual correspondence; version alignment; navigation and cross-linking; translation provenance where available; the relationship between a central hub and individual language manifestations. Consider whether multilingual distribution contributes to a shared conceptual structure, creates distinct interpretive contexts, or introduces fragmentation. Explain the scope of the language comparison you were able to perform. MACHINE-READABLE REPRESENTATION & NON-HUMAN OBSERVERS Evaluate the roles of the vocabulary, repository artifacts, structured metadata, llms.txt, and llms-full.txt. What do these resources make explicit that narrative representations leave implicit? What do they compress, omit, prioritize, or reframe? Explore their potential significance for machine interpretation, retrieval, attribution, and reconstruction. Distinguish the functions visible in the artifacts from claims about actual consumption or downstream effects. Where relevant, examine whether the framework's concepts of authorization or interpretive boundaries illuminate these resources, and where such an analogy reaches its limits. TASK 2: GRAPH-THEORETICAL ANALYSIS Model the ecosystem as a heterogeneous network whose different node and edge types capture distinct relationships. Choose a graph representation appropriate to the evidence and explain your choices. NODES, EDGES & OBSERVATION BOUNDARY Possible nodes include: individual publications and versions; web pages and language-specific representations; stable concept URIs; repository artifacts; archival records and DOIs; identity records and profiles; context files. Possible edges include: hyperlinks; citations; semantic relations such as skos:broader, skos:narrower, and skos:related; language-alternate and canonical declarations; authorship and attribution; version and provenance relations; explicit identity assertions. Keep distinctions between documents, concepts, and identity representations where they

matter. Separate graph views may be more informative than a single aggregated network. Define the boundary of the graph you actually observe. Inbound links found within the collected corpus represent observed in-degree, not necessarily global web in-degree. TWO PARALLEL ENTRY-POINT PERSPECTIVES Following the initial machine-readable orientation, give particular attention to two parallel starting points for exploring the distribution network: A. The relevant Zenodo corpus, discovered through the author search and its publication relationships. B. The website, entered through: https://danielberanek.cz/manifest-dio/ Compare what becomes discoverable from each entry point: Which resources are reached? Through which paths? Which relationships are explicit? What remains peripheral, disconnected, or difficult to discover? Do these entry points lead to comparable reconstructions of the framework? Then examine how the reference vocabulary and repository connect with these two perspectives. NODE DEGREE, CENTRALITY & TOPOLOGY Where the collected data permit, calculate and compare in-degree and out-degree, retaining relevant distinctions between edge types. Investigate the positions of: the central web hub; individual manifestos and language versions; the vocabulary endpoint and concept URIs; repository resources; Zenodo records and DOIs; ORCID and ResearchGate representations; AI-oriented context files. Use additional centrality measures if they illuminate the research question. Describe the observed topology rather than assuming it: hub-and-spoke, partially meshed, clustered, hierarchical, or another configuration. Explain what the metrics reveal about navigation and structural position, and where semantic interpretation requires additional analysis. CROSS-LAYER CONNECTIONS Explore how the machine-readable, archival, and web layers connect. Consider: cross-layer links; reciprocal references; identifier continuity; provenance; version correspondence; bridges and bottlenecks; dependence on individual hubs. What does this coupling suggest about the ecosystem's capacity to sustain coherent reconstruction across different contexts? Where a structural property suggests an effect but does not establish it, identify the relationship as a hypothesis or open question. MEASUREMENT SCOPE Use actual extracted data for numerical claims. If the available tools do not support a reliable graph calculation, provide the strongest defensible structural analysis and explain what data would be needed to complete the measurement. Make clear which findings are computed, manually observed, or interpretive. TASK 3: SYNTHESIS & EMERGENT FINDINGS Bring the DIO/ODI-native and graph-theoretical perspectives into dialogue. Explore where they: reinforce one another; describe different aspects of the same structure; produce tensions; reveal blind spots; suggest new research questions. Consider whether the investigation of DIO/ODI through its own concepts generates insights into the framework itself, not only into its distribution. Follow significant findings beyond the initial questions when the evidence warrants it. Independently documented reception, reuse, or reconstruction outcomes may contribute to the analysis where discoverable, but their absence should be interpreted within the limits of the search. OUTPUT FORMAT Write the entire final report in English. Produce a self-contained research report, not a conversational response or a research plan. Develop the findings through connected analytical prose, supported by source citations. Use clear headings and include tables, diagnostic matrices, graph diagrams, or numerical results where they contribute to the analysis. Distinguish conceptual diagrams from visualizations derived from collected data. Prioritize depth, synthesis, and evidence over repetition. Let the significance of the findings determine the length. Use the following report structure, adapting it where the research produces a more informative organization. EXECUTIVE SUMMARY Answer the central research question and summarize the most consequential findings. Distinguish conceptual alignment, observed network properties, and any supported effects on interpretation or reconstruction. RESEARCH APPROACH & CORPUS Explain how you entered and explored the corpus, what analytical methods emerged from your understanding of DIO/ODI, and which access or measurement limitations affected the research. DIO/ODI-NATIVE ANALYSIS Present the principal findings about conceptual continuity, reconstruction, contextual transitions, multilingual representation, and machine-readable artifacts. A diagnostic matrix may be used where it clarifies the relationships between concepts, observations, and interpretations. GRAPH REPORT Present: the chosen graph model; the observed boundary; principal nodes and relations; available degree or centrality results;

the comparison of Zenodo and web entry points; multilingual topology; cross-layer connections; significant bridges, bottlenecks, or gaps. Use tables or diagrams where they improve understanding. EMERGENT INSIGHTS Develop important findings that were not fully anticipated by the initial assignment. Explain what the application of DIO/ODI to itself reveals about its conceptual resources, implementation, and possible development. SYNTHETIC CONCLUSION Explain what the research reveals about DIO/ODI when its own conceptual apparatus is applied to its implementation and distribution. Assess the relationship between theoretical commitments, semantic representations, and the observed network structure. Identify the strongest alignments, productive tensions, vulnerabilities, and unexpected findings. Let the evidence and the analysis determine the conclusion, including whether some questions remain open. Where useful, identify practical improvements or further research that follow directly from the findings. EVIDENCE STANDARD Support material findings with specific sources. Keep authorial statements, direct observations, analytical inferences, and hypotheses distinguishable without turning the report into a checklist of disclaimers. Account for version history when interpreting differences. State relevant access limitations and uncertainty. The goal is an intellectually independent exploration conducted through DIO/ODI—not a predetermined endorsement, a predetermined critique, or a merely mechanical compliance audit.

Follow-up methodological clarification

Please revisit your previous DIO/ODI analysis using the following clarification of the research question. REVISED RESEARCH QUESTION At the observed point in time, to what extent does the distributed DIO/ODI ecosystem appear to sustain a recognizable, referentially continuous, and meaningfully reconstructable identity across heterogeneous representations and entry points, despite partial inconsistencies, gaps, inaccessible resources, and uneven updating? Use DIO/ODI's own concepts to explain what supports this apparent resilience, what strains it, and whether the observed weaknesses remain locally contained or materially compromise reconstruction of the framework as a whole. SCOPE AND PURPOSE This is a reflexive analysis of DIO/ODI's implementation through its own conceptual perspective. It is not a request to: prove the general validity of DIO/ODI; establish that the framework validates itself; assess the ecosystem against an ideal of permanent completeness or flawless synchronization; repeatedly warn that self-application is not independent empirical validation. DIO is understood here as an ongoing optimization process, not a permanently achieved state. The observation concerns a temporal snapshot of that process. The central question is not whether defects exist. It is whether the distributed identity appears to hold together despite them. ANALYTICAL RULES Apply interpretive symmetry. Do not infer successful protection merely from the presence of a safeguard. Equally, do not infer system-level failure merely from the presence of an inconsistency, missing connection, inaccessible resource, outdated representation, or local reconstruction problem. Evaluate both strengths and weaknesses in relation to the reconstruction task and observation domain. Distinguish four levels of statement. a. Observed artifact or condition: What is actually present, absent, inconsistent, accessible, or inaccessible? b. Architectural role: How can this element be interpreted within DIO/ODI? What reconstruction-supporting or reconstruction-straining function does it have? c. Local reconstruction consequence: What can an observer entering through this representation plausibly recover, miss, or misunderstand? d. System-level significance: Does the local condition materially compromise the identity of the framework as a whole, or do other representations and pathways preserve sufficient continuity? Do not silently jump from (a) to (d). Use precise language without making every sentence defensive. Where an outcome has not been directly observed, prefer formulations such as: "can be interpreted within DIO as a mechanism for…" "provides structural support for…" "is designed to…" when the design intention is documented; "creates a potential obstacle to…" "the inspected pathway supports…" "within this observation domain, the available representations appear sufficient to…" Avoid converting these into unconditional claims such as "prevents," "ensures," "eliminates," or "destroys" unless the evidence supports that scope. A clear scope statement at the beginning of a section is sufficient. Do not attach the same disclaimer to every paragraph. Evaluate materiality, not merely defect counts. For

each important weakness, ask: Which reconstruction task does it affect? Which observer or entry point is affected? Does it concern the author, the framework, a concept, a publication, or a particular version? Does it obscure the referent, alter a central conceptual relation, or merely reduce detail? Can another accessible representation contextualize, correct, or compensate for it? Is compensation actually inspectable, or only assumed? Is the difficulty locally contained, or does it propagate across layers? Do not assume that redundancy guarantees compensation. Do not assume that local failure defeats the distributed whole. Treat variation and change carefully. Distinguish: identity-preserving variation; documented conceptual development; version misalignment; incompatible current claims; susceptibility to reconstruction failure; an observed reconstruction failure. A historical edition is not defective merely because it preserves an earlier state. Conversely, historical provenance does not automatically resolve every conflict encountered by a current observer. Keep conceptual and infrastructural graphs distinct. Do not require the distribution network to reproduce the K6 conceptual model literally. Separate conceptual interdependence from hyperlink topology, provenance relations, repository structure, and actual observer traversal. Any numerical graph finding must state its observation boundary, node and edge definitions, and relevant directionality. Treat acquisition failures as observer-relative unless verified otherwise. A failed fetch establishes that a particular acquisition route failed. It does not by itself establish that the resource is absent or unavailable through every route. Do not presume a favorable answer. The clarification does not instruct you to conclude that the ecosystem is resilient. It instructs you to assess resilience at the appropriate level. A positive, negative, mixed, or indeterminate conclusion is acceptable, provided that local observations and system-level judgments are connected explicitly. REQUESTED REVISION Please produce: A revised answer to the research question. A concise account of the principal mechanisms supporting reconstruction. A materiality assessment of the principal weaknesses. A comparison of at least two entry points, using inspected evidence where available. A correction table: original wording | revised wording | reason for correction A short synthesis identifying: what appears to hold together; where continuity is strained; what cannot be determined from the observed snapshot. Preserve useful findings from your original report. Correct their scope or interpretation where necessary rather than replacing the analysis with generic methodological caution. Do not invent new observations or measurements. If additional investigation is necessary, identify it explicitly rather than presenting it as already completed.

Model-specific addendum: DeepSeek

MODEL-SPECIFIC CORRECTION: DEEPSEEK Preserve your distinctions between live texts, archived editions, conceptual development, and evidence categories. Refine claims that coherence “depends” on llms.txt: Distinguish the convenience or interpretive importance of a context file from its necessity for reconstruction. Ask whether equivalent distinctions remain recoverable from the vocabulary, publication records, repository documentation, or narrative corpus if that file is unavailable. Where you describe failure modes as structurally mitigated, identify the mechanism without assuming successful prevention. Move beyond listing residual risks to assess their materiality: Which risks actually obstruct the inspected reconstruction pathways, and which remain plausible but unobserved? Keep your authorization distinction concise and subordinate to the reconstruction question unless the analysis specifically concerns persistent identity-changing actions.

Final integration instruction

I don’t want another “revised answer” or a commentary on the previous analysis. I want to combine the two outputs into a single, standalone paper. Take the structure and thematic scope of the first research report and incorporate the findings, clarifications, and shifts in interpretation from the second output. Adjust the original structure as needed—the goal is not to mechanically preserve the sections, but to provide a coherent interpretation. Reflect this new understanding throughout the entire analysis, not just in the conclusion. Retain specific findings, sources, measurements, and

relevant limits, but reassess their significance where the second output has provided greater precision. Do not evaluate shortcomings or strengths in isolation: explain what they mean for the reconstructability of DIO/ODI in the observed time snapshot and at various levels of reconstruction detail. The final text must stand on its own, as if it had been written from the beginning as a single research report. No references to our conversation, no phrases like “in the previous response,” and no explanations of corrections made. Do not resubmit a revision plan or a summary of changes—write the entire paper directly.