Review

KG4ESG: The ESG Knowledge Graph Atlas

Chaoyue He *, Xin Zhou, Di Wang, Xinjia Yu, Lei Xiao, Langyue Li, Hong Xu, Wei Liu, Chunyan Miao

Posted Date: 19 March 2026

doi: 10.20944/preprints202602.1970.v2

Keywords: ESG analytics; knowledge graphs; ESG knowledge graphs; large language models (LLMs); sustainability AI; information extraction; entity recognition and linking; relation extraction; ontology engineering; schema design; graph construction; graph representation learning; graph neural networks (GNNs); retrieval-augmented generation (RAG); neuro-symbolic AI; multimodal learning; text-toknowledge graph generation; knowledge graph completion; knowledge graph reasoning; data-toknowledge graph pipelines; knowledge graph applications; ESG data integration; knowledge fusion; provenance-aware systems; sustainability reporting; climate intelligence; carbon accounting; ESG standards; ESG benchmarking; decision support systems; explainable AI (XAI); trustworthy AI; autonomous agents; multi-agent systems

Review

KG4ESG: The ESG Knowledge Graph Atlas

Chaoyue He [1,*], Xin Zhou [1], Di Wang [1], Xinjia Yu [1], Lei Xiao [2], Liangyue Li [2], Hong Xu [1], Wei Liu [2] and Chunyan Miao [1]

[1] Alibaba-NTU Global e-Sustainability CorpLab (ANGEL), Singapore
[2] Alibaba Group, China
* Correspondence: cyhe@ntu.edu.sg

Abstract

Environmental, Social, and Governance (ESG) analytics increasingly uses knowledge graphs (KGs) to encode framework-grounded semantics, align overlapping standards, and attach provenance for auditable querying. Yet ESG evidence is mainly *text-first* (disclosures, regulations, policies, news, incident narratives), so quality depends on the KG–NLP interface. Research remains fragmented across topics, modalities, and pipelines, limiting reuse. To address this, we introduce the **ESG Research Focus Map (ESG-RFM)**, a vendor-agnostic **pillar–theme–focus** taxonomy crosswalked to major ESG frameworks (MSCI, GRI, ESRS, and SASB), which serves as the organizing lens for **KG4ESG**, an atlas-style survey of **337** ESG KG papers (2015–2025). KG4ESG is curated via a query dictionary and PRISMA-style screening across four academic search engines, and provides a structured atlas and reusable resource that organizes the field into two stages: **Data→KG** and **KG→App**. For Data→KG, we summarize evidence sources and distill four construction paradigms: **P1** ontology-first lifting/integration, **P2** rule/supervised NLP extraction, **P3** LLM-assisted structuring/alignment, and **P4** agentic pipelines with iterative validation/repair. For KG→App, we group applications into *reporting & compliance*, *monitoring & risk intelligence*, and *decision support*. A corpus-level meta-analysis highlights gaps in evaluation, openness, and multimodal grounding motivating auditable benchmarks and reusable resources. We will open-source a repository containing all literature, the taxonomy, and related artifacts once permitted.

Keywords: ESG analytics; knowledge graphs; ESG knowledge graphs; large language models (LLMs); sustainability AI; information extraction; entity recognition and linking; relation extraction; ontology engineering; schema design; graph construction; graph representation learning; graph neural networks (GNNs); retrieval-augmented generation (RAG); neuro-symbolic AI; multimodal learning; text-to-knowledge graph generation; knowledge graph completion; knowledge graph reasoning; data-to-knowledge graph pipelines; knowledge graph applications; ESG data integration; knowledge fusion; provenance-aware systems; sustainability reporting; climate intelligence; carbon accounting; ESG standards; ESG benchmarking; decision support systems; explainable AI (XAI); trustworthy AI; autonomous agents; multi-agent systems

1. Introduction

Environmental, Social, and Governance (ESG) regulation, sustainable finance, and corporate accountability increasingly depend on heterogeneous evidence: sustainability reports, regulatory texts with cross-references and exceptions, and narrative incident or controversy records. Because this evidence is largely language-mediated, knowledge graphs (KGs) provide a natural backbone: they encode standards-grounded semantics, represent qualifiers (e.g., time/unit/scope), and attach provenance so that querying and explanation remain auditable.

However, ESG KG research is difficult to synthesize. Works often draw inconsistent topical boundaries, combine partially overlapping standards without explicit alignment, or treat KG construction and language-facing access in isolation. This separation hides a key dependency: what downstream

 doi:10.20944/preprints202602.1970.v2

systems can retrieve, verify, and justify through NLP interfaces is determined upstream by representational choices (schema, qualification, and provenance). In regulatory contexts, the validity of a system's output is contingent upon the auditability of its evidence chain. We therefore frame the field end-to-end (Figure 1) as two coupled stages: **Data→KG** (evidence → schema-governed, provenance-bearing KGs) and **KG→App** (reporting/compliance, monitoring/risk intelligence, decision support) mediated by KG–NLP interfaces such as extraction/linking, text-to-query/KGQA, graph-grounded synthesis, and provenance-aware verification. To make results comparable across topics and standards, we introduce the **ESG Research Focus Map (ESG-RFM)**, a vendor-agnostic pillar–theme–focus taxonomy crosswalked to major frameworks.

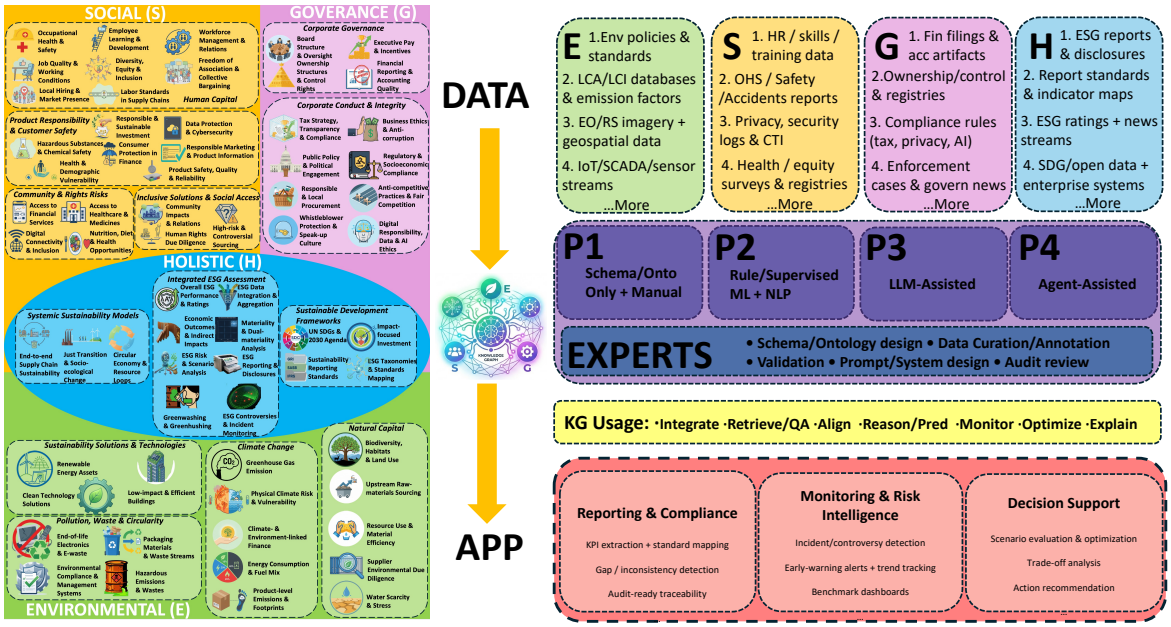

Figure 1. KG4ESG overview. Left: **ESG Research Focus Map (ESG-RFM)** (enlarged in Figure A2); Right: a two-stage pipeline: **Data→KG** construction and **KG→App** utilization, linked through reusable KG–NLP interfaces.

Prior surveys offer valuable perspectives but rarely connect *construction choices* in ESG KGs to *audit-facing behavior* in downstream language interfaces. Large sustainable-finance meta-analyses typically operate at the level of study outcomes, rather than on evidence-linked, machine-readable representations Friede et al. (2015). Ontology surveys in circular-economy and related sustainability domains catalog conceptual artifacts and alignment challenges, but are often decoupled from disclosure- and compliance-oriented pipelines Li et al. (2023). More general roadmaps on integrating KGs with LLMs, RAG, and agentic workflows provide transferable design patterns, yet do not address ESG-specific cross-framework semantics, qualification, and provenance constraints that govern what systems can safely assert Pan et al. (2024); Gao et al. (2023); Guo et al. (2024). Extended illustrations are provided in Appendix A. We additionally position KG4ESG relative to primary technical work on disclosure structuring and cross-standard alignment, compliance- and governance-oriented KGs, and graph-grounded LLM/agent interfaces with validation and provenance in Appendix B.

Contributions. We provide: (1) **ESG-RFM**, a reusable topic map aligned with major ESG standards; (2) the **KG4ESG corpus** of **337** works (2015–2025) retrieved via an ESG-RFM-grounded query dictionary and PRISMA-style screening; (3) a **Data→KG** synthesis that categorizes evidence sources and distills four construction paradigms (P1–P4) together with shifting **expert roles**; (4) a **KG→App** synthesis that organizes applications into reporting/compliance, monitoring/risk intelligence, and decision support, and summarizes recurring KG–NLP interfaces (extraction/linking, KGQA/text-to-query, GraphRAG/RAG-style synthesis, provenance-aware verification); (5) a corpus-level **meta-analysis** of resources, evaluation, and openness signals; and (6) a **future research agenda** grounded in these findings.

 doi:10.20944/preprints202602.1970.v2

2. Survey Methodology

The **KG4ESG corpus** contains **337** papers (2015–2025) whose primary contribution is to *construct, extend, or apply* an ESG/sustainability KG. We follow a taxonomy-driven, PRISMA-style workflow (Appendix D): (1) define **ESG-RFM**, a vendor-agnostic *pillar–theme–focus* map aligned with MSCI/GRI/ESRS/SASB (Appendix C); (2) derive a focus-level **query dictionary** and retrieve candidates from **Google Scholar**, **ACL Anthology**, **OpenAlex**, and **Semantic Scholar**, augmented with backward/forward citation chasing; and (3) merge, deduplicate, and screen records using explicit inclusion/exclusion criteria. Each included work is annotated to connect **Data→KG** choices to **KG→App** behavior through an NLP lens: ESG-RFM topical tags, dominant construction paradigm (P1–P4), evidence modalities and resource statistics, dominant KG–NLP interfaces. Because ESG topics are cross-cutting, we allow multiple focuses but record a single *primary* focus per work.

3. Data-to-KG: Constructing ESG KGs

The **Data→KG** stage turns heterogeneous ESG evidence into a *schema-constrained, provenance-preserving* graph. Across environmental (E), social (S), governance (G), and holistic (H) settings, this stage behaves less like "extracting triples" and more like enforcing an *evidence-to-structure contract*: entity and relation types must be typed, qualified (time, boundary/scope, unit, method), and linked back to supporting evidence so that downstream *language interfaces* can retrieve, explain, and verify claims under shifting standards Diamantini et al. (2022); Yu et al. (2024); Markovic et al. (2023); Yu et al. (2025). The contract is especially non-negotiable in H and G applications (reporting, rating, and compliance) where a statement's meaning depends on definitions and scope, but it is equally decisive in E systems where numeric comparability and spatiotemporal context dominate (e.g., risk, hazards, footprints) Cai et al. (2025); Ershov (2023); Hofmeister et al. (2024); Frakes et al. (2025); Diamantini et al. (2025).

Two representation imperatives recur across all pillars. The first is **qualification completeness**: numeric indicators and event facts must carry units, temporal validity, and boundary/method metadata to remain computable rather than merely retrievable Markovic et al. (2023); Diamantini et al. (2025); Wu et al. (2024); Lippolis et al. (2025). This is most visible in E and H pipelines that compute footprints, LCA metrics, or climate-risk indicators, but it also matters in S contexts such as health, safety, and inclusion metrics that require consistent definitions Nananukul and Kejriwal (2025); Wu et al. (2025). The second is **evidence and provenance linking**, where graph assertions point to clauses, spans, tables, logs, or geospatial products; this turns the KG into a verification surface that survives review, re-labeling, and re-alignment Yu et al. (2025); Ershov (2023); Frakes et al. (2025); Bronzini et al. (2024); Garza et al. (2024).

3.1. Evidence Sources & Modalities

Despite the diversity of domains, most ESG KGs draw from a few recurring evidence families, each with pillar-specific constraints. **Standards, taxonomies, and indicator definitions** provide semantic anchors for H and E work (metric meaning, unit semantics, formulas, cross-framework mapping), enabling comparable computation and alignment across reporting regimes Diamantini et al. (2022); Yu et al. (2024); Diamantini et al. (2025); Zhou et al. (2024). **Regulations, policies, and legal corpora** are central in G and increasingly in H, where obligations, scope, exceptions, and cross-references must be modeled explicitly to support traceable compliance reasoning and review Ershov (2023); Boukhelifa and Merabet (2024); Chung et al. (2025); Tan et al. (2025); Hernandez et al. (2025). **Corporate disclosures and enterprise records** (reports, filings, ERP/MES/CRM tables, process models, event logs) appear throughout, but they are most structurally exploited in G/H reporting and governance workflows, where disclosure structuring, gap detection, and auditable metric management are primary Bronzini et al. (2024); Usmanova and Usbeck (2024); Ong et al. (2025); Osman et al. (2024); Dolha et al. (2025); in S settings, the same "enterprise record" category often shifts toward people- and safety-centered artifacts such as accident/near-miss narratives, dispute or inspection cases, OHS documentation, HR

 doi:10.20944/preprints202602.1970.v2

profiles and mobility histories, resumes and job ads linked to skills taxonomies, and privacy/security documentation (privacy policies, CTI reports, and operational logs), which require tighter context modeling (roles, responsibilities, exposure, and governance constraints) to remain actionable Wu et al. (2025); Liu et al. (2025); Chen et al. (2023); Shi and Wu (2024); Zhang et al. (2025); Fettach et al. (2025); Liu et al. (2025); Cheng et al. (2025); Hu et al. (2024); Cotti et al. (2025); Kurniawan et al. (2024). **External narratives and literature** (news/web/literature) frequently supply H monitoring signals (controversies, trends, reputational risk) and also feed E knowledge acquisition (climate, biodiversity, circular economy) Angioni et al. (2024); Hassan Nassar et al. (2025); Iwata et al. (2025); Islam (2022); Penev et al. (2022); Zhao et al. (2025); they also appear as S-facing evidence when incident reporting and cyber/privacy risk intelligence are constructed from vendor reports, public disclosures, and narrative sources that must be grounded and normalized before use Hu et al. (2024); Zhang et al. (2024); Ren et al. (2022). Finally, **multimodal measurements** (geospatial layers, remote sensing, imagery, sensor/time-series telemetry, CAD/BIM artifacts) are disproportionately important in E and many H systems, making spatiotemporal qualification, multi-resolution linking, and update mechanisms core design goals Hofmeister et al. (2024); Frakes et al. (2025); Zhu et al. (2025); van der Weerdt et al. (2024); Wu et al. (2025); Keena et al. (2025); importantly, these modalities also surface in S deployments where OHS and product-quality settings combine sensor streams and event logs with text to support monitoring, diagnosis, and safety decision support Jiang et al. (2024); Wen (2024); Ren et al. (2024); Wang et al. (2024).

This modality mix yields two recurring construction regimes that often coexist. Text-first pipelines (reports, regulations, news) emphasize *schema-conditioned NLP*: extraction, entity/relation linking, normalization to indicators, and evidence links to clauses/spans, typical in G and H compliance/reporting and in S policy/incident narratives Garza et al. (2024); Zhou et al. (2024); Usmanova and Usbeck (2024); Iwata et al. (2025). Sensor/geo-first pipelines emphasize integration, spatiotemporal semantics, and provenance over derived products, typical in E monitoring (hazards, infrastructure resilience, water/energy) and in H disparity/impact analytics that blend geo-data with socio-economic context Hofmeister et al. (2024); Frakes et al. (2025); Lippolis et al. (2025); van der Weerdt et al. (2024). Hybrid designs increasingly bind narrative evidence to geospatial and imagery context for disaster and climate-risk analysis, or bind built-environment models to IoT streams for multi-scale sustainability evaluation Wu et al. (2025); Aivalis et al. (2025); Chen et al. (2025); Quek et al. (2024); Karjou et al. (2025). Thus, the *text-to-typed-structure* problem dominates ESG KGs.

3.2. Paradigms (P1–P4) & Pillar Pressures

Four recurring construction paradigms appear across E/S/G/H. They differ in dominant evidence-to-schema mapping operator, but all rely on explicit schema conformance, constraint checking, and provenance-aware validation to prevent semantic drift and to preserve auditability Diamantini et al. (2022); Yu et al. (2024, 2025)(Appendix E).

3.2.1. P1: Ontology-First & Manual Integration

P1 approaches prioritize ontology stewardship and deterministic mappings. They are common in E and H contexts where interoperability and computability are non-negotiable: emissions-factor semantics and provenance of calculations Markovic et al. (2023), LCA indicator computation and cross-organization comparability Diamantini et al. (2022, 2025); Saad et al. (2023), and large-scale geo-semantic integration over heterogeneous datasets Frakes et al. (2025); Zhu et al. (2025); Wu et al. (2022). They also appear in E/H built-environment and IoT settings where reused vocabularies enable consistent linking of devices, spaces, and measurements van der Weerdt et al. (2024); Karjou et al. (2025); Delgoshaei et al. (2022). The strength is replayable semantics and clean computation; the trade-off is that coverage and agility depend on curated mappings and disciplined schema evolution as standards and data sources change Diamantini et al. (2022); Markovic et al. (2023).

 doi:10.20944/preprints202602.1970.v2

3.2.2. P2: Rule/Supervised NLP/ML Extraction

P2 pipelines treat the schema as a label space populated by supervised NLP, patterns, and hybrid rules. This is prominent in S pipelines that structure incidents and occupational safety narratives for causal or risk analysis Chen et al. (2023); Shi and Wu (2024); Liu and Yang (2022); Liu and Cheng (2024); Li et al., and in security/privacy-adjacent S applications where structured representations are extracted from reports and logs Ren et al. (2022); Shen et al. (2020); Yin et al. (2024). In G and H settings, P2-style extraction also appears in filings/news monitoring and in fraud/risk corpora when the target labels are stable Angioni et al. (2024); Oksanen et al. (2022); Shen et al. (2021); Wu et al. (2022). P2 scales ingestion effectively, but ESG concept drift (new reporting requirements, new controversy categories, evolving definitions) increases maintenance costs, motivating alignment-aware and validation-first extensions when target semantics move Diamantini et al. (2022); Angioni et al. (2024).

3.2.3. P3: LLM-Assisted Construction

P3 pipelines use LLMs to produce schema-conditioned structured outputs, typically coupled with retrieval grounding and validation. This has become especially influential in H and G systems that must structure sustainability reports, align indicators across frameworks, and maintain traceable links to definitions and evidence Yu et al. (2025); Bronzini et al. (2024); Zhou et al. (2024); Usmanova and Usbeck (2024). Similar patterns increasingly support S domains (policy, health, safety guidance) where heterogeneous narratives and long-tail terminology make brittle rules less attractive Wu et al. (2025); Zhang et al. (2025). Because LLM outputs can introduce silent schema drift, robust P3 systems emphasize provenance preservation and explicit validation hooks (type/unit checks, constraint verification, evidence links), framing evaluation at both fact level and schema/ontology level Yu et al. (2025); Cai et al. (2025); Garza et al. (2024).

3.2.4. P4: Agentic/Tool-Using Pipelines

P4 treats construction as a controlled retrieve–extract–validate–repair loop, integrating query engines, calculators, geospatial operators, optimization tools, and producing replayable traces for audit and continual updates Hofmeister et al. (2024); Quek et al. (2024); Jomraj et al. (2025). This is common in G/H audit-facing apps (traceable compliance QA, evidence-backed explanations) Tan et al. (2025); Jomraj et al. (2025), in H benchmarking-oriented ESG extraction and scoring where quality criteria are enforced iteratively Cai et al. (2025), and in E operational sustainability analytics where completion, enrichment, and emissions accounting couple KG completion with optimization and scenario comparison Diamantini et al. (2025); Felder et al. (2025); Hou et al. (2025). Here, trust is externalized into constraints and traces, turning KG into a governed process artifact rather than a one-shot output Cai et al. (2025); Jomraj et al. (2025).

3.3. Cross-Cutting Representation Choices

Three representation choices consistently determine downstream feasibility. **Unit and quantity modeling** is essential once the KG must support numeric retrieval and indicator computation, particularly in E/H footprints and operational carbon traces, but also in S/H equity and health metrics where comparability depends on explicit definitions Markovic et al. (2023); Diamantini et al. (2025); Wu et al. (2024); Nananukul and Kejriwal (2025). **Spatiotemporal modeling** becomes mandatory in E and many H settings that integrate sensor streams, geospatial layers, and event logs, enabling multi-resolution analysis for hazards, infrastructure resilience, and disparity assessment Hofmeister et al. (2024); Frakes et al. (2025); Lippolis et al. (2025); van der Weerdt et al. (2024). **Provenance modeling** bridges extraction and audit across all pillars: it supports inspection, re-evaluation, and dispute resolution as standards, contexts, and evidence quality shift Yu et al. (2025); Ershov (2023); Bronzini et al. (2024); Garza et al. (2024).

 doi:10.20944/preprints202602.1970.v2

3.4. How Construction Shifted Over Time

Early systems (roughly prior to 2021) emphasized interoperability and curated semantics, favoring ontology-first integration in E and early cross-domain settings Yan et al. (2018); Elluri et al. (2018); Atzeni et al. (2020). As monitoring and extraction expanded (2021–2023), supervised and rule-based pipelines grew for S incident/safety corpora and for G/H disclosure and risk corpora Liu and Yang (2022); Shen et al. (2020); Oksanen et al. (2022); Mishra and Mittal (2021). From 2022 onward, rapid growth in standards and cross-framework pressure pushed H and G systems toward flexible alignment and long-tail coverage, accelerating LLM-assisted structuring with explicit validation Diamantini et al. (2022); Yu et al. (2025); Zhou et al. (2024); Usmanova and Usbeck (2024). In parallel, auditability and streaming updates in E/H settings strengthened iterative, validation-first workflows that expose traces and constraints as part of the construction contract Cai et al. (2025); Quek et al. (2024); Jomraj et al. (2025).

3.5. Quality Control & Expert Roles

Expert oversight remains central, but its leverage point also shifts. In E and H, experts often shape ontologies and calculation semantics (units, boundaries, conversion factors, indicator formulas) Diamantini et al. (2022); Markovic et al. (2023); Diamantini et al. (2025). In S settings, expert labor frequently concentrates on annotation and validation of incident factors, causal chains, and safety/health categories Liu et al. (2025); Shi and Wu (2024); Liu and Yang (2022). In G and H compliance, expertise concentrates on encoding scope/exceptions, validating obligations, and governing acceptance via constraints and audit procedures Ershov (2023); Chung et al. (2025); Tan et al. (2025); Jomraj et al. (2025). Across pillars, the same reliability hooks recur: qualification completeness (time, unit, scope, method) for KPIs and events Markovic et al. (2023); Wu et al. (2024); Lippolis et al. (2025); standards-to-KG traceability for indicator mapping and disclosure alignment Diamantini et al. (2022); Yu et al. (2024); Zhou et al. (2024); evidence links to clauses, spans, tables, logs, and geospatial products Ershov (2023); Frakes et al. (2025); Garza et al. (2024); constraint enforcement and quality reporting Yu et al. (2025); Cai et al. (2025); and reproducible update traces for longitudinal monitoring and audit Quek et al. (2024); Jomraj et al. (2025); Vanapalli et al. (2025).

4. KG-to-App: ESG Applications

The **KG→App** stage operationalizes ESG KGs as structured backbones for reporting and compliance, monitoring and risk intelligence, and decision support across pillars. Here the KG acts simultaneously as a **structured index** over heterogeneous evidence and a **reasoning substrate** whose schema, qualifiers, and provenance make outputs inspectable and auditable Diamantini et al. (2022); Markovic et al. (2023); Ershov (2023). Crucially, ESG systems are typically interacted with through *language*: users ask questions, reviewers demand justifications, and evidence is authored in text. Thus the dominant integration point is the KG–NLP interface: controlled retrieval (often over evidence-linked subgraphs), text-to-query/KGQA, graph-grounded context expansion (GraphRAG/RAG), and provenance-aware verification Cai et al. (2025); Tan et al. (2025); Jomraj et al. (2025). This is why provenance-aware retrieval and graph-grounded context construction are increasingly used as the default interface for narrative-heavy evidence, particularly in H/G reporting, compliance, and monitoring of controversies and reputational risk Garza et al. (2024); Ushio et al. (2025); DeBellis et al. (2025). (Appendix F)

4.1. Interaction Patterns Across E/S/G/H

Four interaction patterns recur across pillars. First, the KG enables **evidence grounding**, binding answers and scores to explicit evidence so reviewers can verify what was used; this is especially central in G/H compliance and in H claim verification Garza et al. (2024); Tan et al. (2025); Jomraj et al. (2025); Kaoukis et al. (2025). Second, it supports **querying and KGQA**, including natural-language access for relational exploration and structured search; this appears in E/H administrative and regulatory

 doi:10.20944/preprints202602.1970.v2

settings and in S safety/health QA where the value is controlled retrieval over structured relations Tan et al. (2025); Zhang et al. (2025); Ushio et al. (2025); Wang et al. (2025). Third, it enables **KG-enhanced learning and prediction**, where graph features and reasoning paths support interpretable modeling of fraud, risk propagation, controversies, and operational risks, spanning G financial governance, H ESG intelligence, and S safety/security settings Iwata et al. (2025); Shen et al. (2021); Wu et al. (2022); Cai and Xie (2024); Wang et al. (2025); Zhu et al. (2025). Fourth, it supports **optimization and scenario evaluation**, where entities, constraints, and KPIs define a world model for planning and benchmarking, common in E circular economy and sustainable logistics, and increasingly used in H policy/SDG planning scenarios Diamantini et al. (2025); Felder et al. (2025); Hou et al. (2025); Benjira et al. (2025).

A distinctive ESG-wide requirement is **numeric indicator computation under semantic constraints**. When the output is a KPI (emissions, footprint, risk score), the system must reconcile units, methods, and boundaries rather than merely retrieving numbers, which is most visible in E/H footprinting and climate-risk analytics but also applies to S/H equity and health indicators Markovic et al. (2023); Diamantini et al. (2025); Wu et al. (2024); Nananukul and Kejriwal (2025). This drives tight coupling between KG semantics and application logic in lifecycle, circular-economy, and infrastructure contexts where constraint-aware planning and comparable computation are core objectives Hofmeister et al. (2024); Zhao et al. (2025); Lu et al. (2024); Yin et al. (2025).

4.2. Application Families with Pillar Emphasis

Reporting and compliance applications lean heavily on H and G semantics: they structure disclosures, align indicators across standards, compute or normalize metrics, and detect gaps or inconsistencies between claims and requirements Diamantini et al. (2022); Yu et al. (2025); Bronzini et al. (2024); Zhou et al. (2024); Usmanova and Usbeck (2024); Ong et al. (2025). Legal and governance deployments emphasize clause-level reasoning with cross-references and scope, enabling traceable compliance checking and audit-ready review Ershov (2023); Boukhelifa and Merabet (2024); Chung et al. (2025); Tan et al. (2025). Parallel patterns appear in privacy, AI governance, and tax settings, where obligations are represented structurally and outputs are backed by evidence-linked access paths designed for audit Garza et al. (2024); Tan et al. (2025); Hernandez et al. (2025); Tauqeer et al. (2022); Zhang (2025). As these systems become audit-facing by default, they increasingly return provenance-grounded traces and inspectable supporting subgraphs so judgments remain verifiable end-to-end Tan et al. (2025); Jomraj et al. (2025).

Monitoring and risk-intelligence applications are primarily H, but they draw signals from E, S, and G sources. They include controversy and violation detection from news and web streams Angioni et al. (2024); Hassan Nassar et al. (2025); Iwata et al. (2025), event-centric S safety analytics from accident and near-miss corpora Liu et al. (2025); Chen et al. (2023); Shi and Wu (2024); Liu and Yang (2022); Zhang et al. (2025), and supply-chain transparency/risk monitoring where graph completion and path reasoning surface hidden dependencies and support evidence-linked risk narratives Kosasih et al. (2024); AlMahri et al. (2025); Jin et al. (2025); Heus et al. (2025); Zheng et al. (2025). E-facing monitoring extends to geo/sensor domains such as outage disparity analysis and multi-resolution spatiotemporal reasoning Frakes et al. (2025), and to disaster/climate-risk representations that combine narrative reports with geospatial and imagery evidence for interpretable comparison and decision support Hofmeister et al. (2024); Aivalis et al. (2025); Chen et al. (2025); Marotta et al. (2024). Where misinformation, greenwashing, or claim manipulation is central (often H), claim-verification and fact-checking graphs provide provenance-linked structures for transparent retrieval and explanation DeBellis et al. (2025); Kaoukis et al. (2025); Burel and Alani (2025).

Decision-support apps are strongest in E and S but regularly couple into G/H governance requirements. E KGs act as constraint-aware substrates for optimization and scenario evaluation in sustainable logistics and circular economy planning Zhao et al. (2025); Felder et al. (2025); Hou et al. (2025); Zhao et al. (2025). Infrastructure and process contexts (water/energy/buildings) integrate telemetry, process knowledge, standards, and enterprise records to support querying, diagnosis, and operational planning Lippolis et al. (2025); Wu et al. (2025); Quek et al. (2024); Karjou et al. (2025);

 doi:10.20944/preprints202602.1970.v2

Papageorgiou et al. (2023, 2025). Product-quality and lifecycle settings combine S safety expectations with operational traceability, supporting end-to-end tracing and feedback control Ren et al. (2024); Jing and Li (2024); Zang et al. (2025), and disassembly/recycling planning under dependency constraints Yin et al. (2025); Wu et al. (2024); Wang et al. (2023). Social-oriented decision support spans nutrition/health guidance and equity analytics (S/H), as well as skills and education pathway recommendation (S) Nananukul and Kejriwal (2025); Wu et al. (2025); Fettach et al. (2025); Fu et al. (2023); Gao et al. (2025); Weichselbraun et al. (2022); Fathi (2024).

4.3. Cross-Cutting Tasks, Scale, & Evolution

Three task clusters recur across application families. One is **benchmarking and scoring with explanations**, including H ESG evaluation and G fraud/risk modeling where explanations must be tied to structured evidence and interpretable factors Cai et al. (2025); Wu et al. (2022); Cai and Xie (2024); Cai et al. (2025). Another is **SDG and sustainability analytics** (H) that compute indicators and analyze interlinkages from heterogeneous open data with provenance and conflict resolution Benjira et al. (2025); Fotopoulou et al. (2022); Androna et al. (2024). The third is **audit-facing interaction contracts** (especially G/H) requiring evidence links and reproducible traces rather than narrative-only outputs Ershov (2023); Tan et al. (2025); Jomraj et al. (2025).

Scale strongly shapes app design and tends to correlate with pillar data modalities. Document- or case-level systems are common in G/H compliance and standards alignment, where controlled review is central Yu et al. (2025); Zhou et al. (2024); Jomraj et al. (2025). Very large-scale graphs more often occur in E/H geo/sensor integration and monitoring, where update strategies, multi-resolution reasoning, and performance constraints dominate Frakes et al. (2025); Lippolis et al. (2025); Zhu et al. (2025); van der Weerdt et al. (2024). Historically, early work often treated KGs primarily as interoperability and querying layers over curated semantics Yan et al. (2018); Elluri et al. (2018); Atzeni et al. (2020); later, S incident/safety corpora and G/H disclosure/news monitoring expanded the risk-intelligence surface Angioni et al. (2024); Liu and Yang (2022); Oksanen et al. (2022); and more recently, the rapid expansion of standards and streaming signals increased demand for cross-framework alignment and auditable, validation-first interactions where traceability is an explicit output rather than an afterthought Yu et al. (2025); Cai et al. (2025); Zhou et al. (2024); Usmanova and Usbeck (2024); Jomraj et al. (2025); Vanapalli et al. (2025).

5. Meta-Analysis

(i) Trend signal from query volumes. The query-volume analysis shows monotonic KG uptake from 2015–2025 with a clear acceleration after 2021 across engines, while KG-based work remains a small fraction of the broader ESG literature (Figure A11, Figure A12, Figure A13). At the focus level, KG activity consistently concentrates in *Data Protection & Cybersecurity*, *Energy Consumption & Fuel Mix*, and *Digital Responsibility, Data & AI Ethics*, alongside core climate indicators (GHG emissions, product footprints) and SDG analytics (Figure A7, Figure A9, Figure A10). Notably, governance-heavy foci such as *Digital Responsibility, Data & AI Ethics* exhibit among the highest KG adoption rates in high-volume topics (e.g., exceeding 5% on Google Scholar), while the ACL Anthology shows substantially higher per-focus KG rates (often 10–20%; Figure A8), indicating that KG4ESG progress is amplified through *NLP-facing interfaces* such as schema-conditioned extraction, alignment, KGQA/text-to-query, and evidence-grounded verification. **(ii) Resource–application mismatch.** A recurring gap is misalignment between **Data**→**KG** and **KG**→**App** contributions: construction papers introduce schemas or pipelines without releasing machine-readable artifacts, while application papers often rely on proprietary graphs or undisclosed mappings; we recommend reporting an explicit contribution footprint per work (schema + provenance model + validation rules + KG access; or task protocol + evaluation + explicit output→evidence linkage), aligned to pipeline stage. **(iii) Openness remains uneven.** Fully reusable releases (schema *and* non-trivial KG access) remain a minority, limiting reproducibility and benchmark formation; a lightweight "openness profile" (schema, data, code, endpoint) should become standard, and a small set of *fully open* ESG KGs spanning multiple pillars would anchor shared evaluation. **(iv)**

 doi:10.20944/preprints202602.1970.v2

Shift toward validation-first pipelines. LLM- and tool-driven workflows increasingly *wrap* (rather than replace) ontology- and supervised components, shifting technical risk to grounding, constraint enforcement, provenance integrity, and audit trails; consequently, **validation-first design** (typed outputs, unit and scope checks, evidence links, conflict reporting) is becoming the default for high-stakes ESG use. **(v) Modality skew and topical long tail.** E/H systems more often integrate geospatial or sensor evidence, while S/G work is dominated by text and tabular sources, leaving multimodal grounding under-explored; focus-level activity is uneven, with infrastructure- and governance-heavy foci (cybersecurity, AI ethics, energy and emissions) leading KG uptake, while many sensitive S/G foci remain under-served. Details are in Appendix G.

6. Future Research Agenda

Given post-2021 acceleration and sustained topical pull from cybersecurity, digital responsibility/AI ethics, and climate–energy indicators (Figure A11–Figure A13), the next breakthroughs are likely to come from **auditable end-to-end contracts** rather than isolated components: cross-framework normalization that preserves units/scope, compliance-grade KGQA/verification that returns evidence-linked subgraphs and traces, and evaluation that scores extraction/alignment/grounded generation under explicit schema and provenance constraints.

(1) Data→KG benchmarks for extraction, normalization, and alignment. Shared benchmarks should test schema-constrained extraction and linking with qualification (time, unit, scope, method) and provenance completeness, and explicitly evaluate cross-standard mappings (GRI–ESRS–SASB–MSCI–SDGs) while penalizing silent schema drift. Recent ESG LLM evaluation benchmarks (e.g., ESGenius) provide complementary task scaffolds that could be extended with KG-grounded qualification scoring and schema-conformance evaluation He et al. (2025). **(2) KG→App task suites with auditable reasoning.** Downstream evaluation needs shared task suites (compliance QA, disclosure gap detection, claim verification, due diligence, risk monitoring) whose protocols require *KG-level evidence* (nodes, edges, provenance) alongside answers. **(3) Expert-in-the-loop and stakeholder-aware workflows.** Expert roles (schema design, annotation, validation, governance) should be reported and linked to provenance and versioning; research should measure the marginal value of expert feedback in P3–P4 pipelines and design escalation paths for sensitive topics. **(4) Closing the text–evidence gap with multimodal ESG KGs.** Progress requires KGs that unify text, tables, sensor streams, and geospatial layers with consistent entity and event models and cross-modal linking, enabling "show the evidence" behavior for both narrative and numeric outputs. This need is consistent with emerging multimodal ESG reasoning benchmarks such as MMESGBench, which further motivate unified cross-modal grounding and evaluation protocols Zhang et al. (2025). **(5) Green NLP and sustainable KG pipelines as operational constraints.** Future work should report compute and energy cost for construction and querying, and develop incremental update strategies that avoid full rebuilds; provenance should record pipeline stages and model or tool configurations to make efficiency–quality trade-offs comparable. **(6) Community reuse and extension of ESG-RFM.** ESG-RFM should evolve as a versioned community artifact with stable identifiers and maintainable crosswalks to evolving standards; reusing it to tag datasets, benchmarks, and systems can reduce fragmentation and make coverage comparable across the ecosystem. **(7) Illuminate ESG-RFM blind spots.** Across pillars, several high-stakes foci remain under-studied in KG4ESG and warrant explicit targeting: `SuppEnvDD`; `FoACollB`; `LocHirMarPre`; `FinProt`; `RespMkt`; `BoardOvr`; `ExecPay`; `PolEngage`; `RespProc`; `Whistle`; and `ImpInv`. Addressing these gaps likely requires more efforts and attention in the future.

7. Conclusion

KG4ESG provides ESG-RFM and a curated corpus of 337 ESG KG works (2015–2025), organized as **Data→KG** construction and **KG→App** use. Query-volume trends show accelerating KG uptake after 2021, led by compliance- and governance-constrained foci alongside core climate/energy indicators. Across domains, ESG KGs are increasingly operationalized through *language interfaces*—which makes

 doi:10.20944/preprints202602.1970.v2

qualification and evidence linkage first-class constraints. These patterns motivate auditable, validation-first pipelines and shared benchmarks that require evidence-linked outputs.

Limitations

Our survey is constrained by the scope of the KG4ESG corpus and the search strategy used to construct it. The 337 papers we analyze are the result of a PRISMA-style pipeline driven by a query dictionary anchored on the phrase "knowledge graph", combined with manual venue search and citation chasing. This setup privileges work that explicitly self-identifies as KG- or ontology-based and is discoverable through mainstream search engines, which means that adjacent literature framed in terms of databases, graphs, or ontologies but not using the "knowledge graph" label may be under-represented. Likewise, research published in non-indexed venues, paywalled reports, and grey literature is only partially captured. Although we attempted to harmonize coverage across environmental, social, governance, and holistic topics, the resulting corpus should be interpreted as a carefully curated but incomplete sample rather than an exhaustive census of ESG KGs.

The taxonomy and coding scheme we introduce are also subject to design choices and subjective judgment. Mapping multiple industrial standards (ESRS, GRI, SASB, MSCI) into a single pillar–theme–focus hierarchy necessarily involves simplification and normative decisions about which distinctions to preserve and which to merge. Similarly, assigning each paper to a primary pillar, theme, and focus area, as well as to a single dominant construction paradigm (P1–P4), abstracts over the fact that many systems span multiple topics and methodological families. We do not assume focus areas are mutually exclusive: ESG phenomena, standards, and applications are structurally cross-cutting (e.g., transition planning couples climate metrics with governance controls and workforce impacts; biodiversity overlaps with land use and water; privacy and cybersecurity often intersect both social and governance concerns). Enforcing strict exclusivity would either duplicate papers across categories or require arbitrary splitting that obscures real dependencies. We therefore allow multi-focus tagging, and use a single primary focus only for aggregation; as a result, focus-level plots should be interpreted as "dominant emphasis" rather than "exclusive membership." The released annotations are designed to preserve these overlaps explicitly so that readers can regroup or re-aggregate the corpus under alternative (but still valid) interpretations. Quantitative summaries and visualizations should therefore be read as indicative patterns rather than precise measurements. In addition, ESG-RFM is primarily a **topic-level research map** that we use consistently across the survey: it structures corpus retrieval (via the query dictionary), supports comparable coding of both Data→KG and KG→App papers, and provides a shared index for summarizing how data sources, KG artifacts, and downstream systems distribute across ESG topics. At the same time, some infrastructure- or standards-centric contributions do not naturally align with a single focus area; these papers are therefore best interpreted through the construction paradigm and released-artifact lens rather than through a single topical assignment.

The temporal analysis is limited by both the collection window and the fast-moving nature of LLM- and agentic methods. Our search was conducted in late 2025 and focuses on work published between 2015 and 2025. New KGs, paradigms, and benchmarks are likely to appear after this window, and some very recent preprints or industrial systems may not yet have been captured at the time of coding. In addition, we rely on publication year as a coarse proxy for methodological evolution, even though development and deployment timelines can lag behind publication, and older systems may have been retrofitted with LLM- or agentic components that are not described in the original papers.

Methodologically, our meta-analysis is based entirely on reported descriptions; we do not re-implement systems, re-run extraction pipelines, or independently audit evaluation setups. Claims about performance, scalability, or robustness are therefore taken from the original publications and may not be directly comparable across domains and tasks. The same caveat applies to openness and resource availability: many KGs are only partially released (e.g., schema but not data, or code without underlying corpora), and in some cases we were unable to verify whether promised artifacts are still

 doi:10.20944/preprints202602.1970.v2

hosted or maintained. We will open-source the taxonomy, query dictionary, annotations, and analysis code once permitted.

Finally, the perspective of this survey is deliberately centered on NLP- and KG-driven views of ESG and sustainability. We focus on text-centric pipelines and multimodal integrations where KGs are explicit artifacts, and do not attempt to synthesize the much broader literature in climate science, ecology, public policy, economics, or critical social sciences that study ESG topics without using KG terminology. As a result, some substantive debates about the meaning and validity of ESG metrics, materiality, or just transition appear here only indirectly, through how they are encoded in KGs. Our identification of "gaps" and "opportunities" reflects this KG4ESG lens and should not be read as a prescriptive statement about what ESG research as a whole ought to prioritize.

Ethical Considerations

This work is a secondary analysis of published research on ESG and sustainability KGs. We do not collect new human-subject data, conduct experiments with users, or work with personally identifiable information beyond what is already present in scientific publications and bibliographic metadata. Our corpus consists of papers that have themselves passed peer review and, where applicable, institutional ethics procedures. Nevertheless, by systematizing this literature and proposing a taxonomy and meta-analysis, we can influence how KG4ESG technologies are developed and deployed, and it is important to reflect on the associated ethical dimensions and potential risks.

A first concern is representation and epistemic fairness. Our search strategy, reliance on English-language venues, and focus on widely indexed conferences and journals biases the corpus toward systems built by institutions in the Global North and toward domains with strong digital or regulatory infrastructures. KGs and applications originating from under-represented regions, languages, and communities are likely under-sampled, and topics that are salient in Global South contexts (e.g., informal economies, customary land tenure, community-led monitoring) may be under-exposed. By releasing our taxonomy and corpus annotations, we hope to make these biases visible and enable other researchers to augment or contest our coverage, but this survey does not resolve underlying structural imbalances in who produces ESG data and knowledge.

Second, many of the KGs we survey operate in high-stakes domains such as credit and insurance, taxation, health equity, labor management, and compliance. Techniques for ESG scoring, risk modeling, and supply-chain visibility can be used to advance sustainability and social protection, but they can also enable more granular worker surveillance, exclusionary lending, or the externalization of risk to vulnerable communities. Similarly, graph-based and LLM-assisted ESG analytics can lend an unwarranted air of objectivity to indicators that remain politically contested. Our role in this survey is descriptive rather than prescriptive: we summarize systems and highlight emerging capabilities, but do not endorse specific ESG metrics or rating practices. We encourage practitioners to use KG4ESG methods within robust governance frameworks that include domain experts, affected stakeholders, and mechanisms for appeal and redress.

Third, although KG4ESG is a survey, the accompanying artifacts we plan to release (ESG-RFM, the query dictionary, and corpus-level annotations/metadata) may lower the barrier to building ESG monitoring, scoring, and compliance systems. This creates *dual-use* and misuse risks: (i) automated ESG screening in hiring, lending, insurance, procurement, or supply-chain due diligence could amplify biases present in source evidence and lead to unfair exclusion; (ii) fine-grained monitoring pipelines may enable intrusive worker, supplier, or community surveillance under the guise of sustainability oversight; (iii) controversy- and incident-centric graphs sourced from news or narrative reports can propagate unverified allegations and cause reputational harm if treated as ground truth; and (iv) cross-framework crosswalks and graph-backed explanations can project a misleading sense of certainty ("false authority"), even when standards are ambiguous, evolving, or contested. We therefore frame KG4ESG outputs as *decision support* rather than automated adjudication, and we encourage downstream users to (a) require provenance-linked evidence and explicit uncertainty/confidence

 doi:10.20944/preprints202602.1970.v2

reporting, (b) keep humans in the loop for high-stakes determinations, (c) perform bias and harm assessments with domain experts and (where feasible) affected stakeholders, and (d) apply appropriate access controls, security review, and governance when KGs incorporate proprietary or personal data. To further reduce harm from errors, we will version and time-stamp released artifacts and maintain a mechanism for reporting and correcting problematic mappings or annotations.

Fourth, there is a risk that the tools and architectures surveyed here contribute to more sophisticated forms of greenwashing or "sustainability-washing". KGs can be used to curate narratives of alignment with standards, taxonomies, and SDGs without a commensurate change in underlying practices, and LLM-based interfaces may make it easier to generate plausible but misleading sustainability stories. Our discussion of future research emphasizes provenance, evidence grounding, and links to physical and regulatory data precisely because these can support external scrutiny and falsifiability, but implementing such safeguards ultimately depends on how organizations choose to use the technologies we describe. Because ESG KGs span both Data→KG (what is encoded and how it is validated) and KG→App (how it is queried and acted upon), governance risks arise end-to-end: weak provenance or schema drift in construction can propagate into downstream decisions, while persuasive language interfaces can obscure uncertainty unless outputs are explicitly linked to KG evidence and versioned artifacts.

Fifth, the environmental footprint of KG4ESG pipelines is itself non-negligible. While many ontology-first and rule-based systems are relatively lightweight, the recent shift toward large neural models, multimodal fusion, and agentic workflows implies substantial compute and energy use. As a survey, we do not estimate the carbon footprint of individual KGs or models, and many of the primary papers do not report energy metrics. Nonetheless, we see it as ethically important that future KG4ESG work reports training and inference costs where feasible, considers lower-footprint alternatives, and avoids deploying unnecessarily heavy models in settings where simpler methods suffice.

Sixth, some of the KGs in our corpus are built on sensitive or proprietary data (e.g., financial statements, tax records, HR databases, clinical or claims data). We interact with these systems only through their published descriptions and do not attempt to reconstruct or redistribute underlying datasets. When we release our own artifacts, we restrict ourselves to bibliographic and high-level metadata about the papers and systems, without including any personal data beyond what is already public in author lists and affiliations. We encourage future work building on our resources to follow privacy- and security-preserving best practices, including data minimization, appropriate access controls, and attention to the risks faced by individuals and communities represented in ESG KGs. We also avoid redistributing copyrighted full texts or any underlying proprietary corpora, and instead release only derived metadata and annotations needed for reproducibility.

Potential Systemic and Deployment Risks. Finally, we highlight the broader potential risks introduced by the widespread operationalization of ESG KGs in automated financial and regulatory ecosystems. As investment algorithms, credit rating agencies, and supply-chain platforms increasingly rely on shared or similarly constructed graph representations, the ecosystem faces the risk of *algorithmic herding*. A systematic bias, mapping error, or hallucinated alignment in a widely adopted ESG KG could cause multiple institutions to simultaneously misprice assets, unfairly penalize specific sectors, or withdraw capital from vulnerable developing regions. Furthermore, the integration of agentic pipelines (P4) introduces *automation bias*, where human auditors and policymakers may defer to the graph's structured, evidence-linked outputs as absolute truth, overlooking nuanced contextual realities or qualitative factors that the schema failed to capture. Mitigating these systemic risks requires treating ESG KGs not as objective oracles, but as inherently lossy, opinionated models that demand continuous stress-testing, diverse human oversight, and transparent fail-safe mechanisms when integrated into high-stakes automated workflows.

Acknowledgments: This research is supported by the RIE2025 Industry Alignment Fund (Award I2301E0026) and the Alibaba-NTU Global e-Sustainability CorpLab.

 doi:10.20944/preprints202602.1970.v2

Appendix A. Extended Discussion of Prior Surveys

This section expands the positioning of prior survey strands briefly referenced in Section 1. It does not introduce new notation or assumptions.

Appendix A.1. ESG & Sustainable Finance Surveys

A major strand of prior work in sustainable finance consists of large-scale reviews and meta-analyses that synthesize evidence on the association between environmental, social, and governance (ESG) criteria and corporate financial performance across a very large body of empirical studies Friede et al. (2015). These reviews organize findings by study design (e.g., portfolio vs. non-portfolio studies), regions, and asset classes, and they highlight that results can vary with methodological choices and context Friede et al. (2015). From the perspective of KG4ESG, the key limitation is that this literature is not primarily concerned with producing explicit, machine-readable representations of the underlying evidence (standards, disclosures, policies, and narratives); instead, it aggregates conclusions reported in the finance literature. KG4ESG is complementary in that it targets evidence-to-structure representations that can be queried, aligned across standards, and audited.

Appendix A.2. Sustainability Ontologies & Domain KGs

Surveys of sustainability- and circular-economy-related ontologies provide a complementary perspective by cataloguing conceptual artifacts for representing cross-industry sustainability domains. For example, Li et al. (2023) surveys general ontologies relevant to the circular-economy domain and related cross-industry areas (e.g., sustainability, materials, manufacturing, products, and logistics), and discusses interoperability challenges that arise when reusing and aligning heterogeneous ontologies. Such work underscores the importance of explicit schemas and alignment strategies when integrating sustainability data.

Appendix A.3. KG Construction, LLMs, RAG, & Agentic Workflows

General KG surveys provide methodological context for KG4ESG. Hogan et al. (2021) offers a broad introduction to knowledge graphs, including data models, querying and validation, and deductive/inductive techniques for making knowledge explicit. Complementary surveys review KG representation learning, acquisition and completion, temporal KGs, and knowledge-aware applications Ji et al. (2021). In the LLM era, recent work also evaluates how LLMs can be used for KG construction and reasoning tasks (e.g., extraction, link prediction, and KGQA) and discusses future opportunities, including multi-agent approaches for KG workflows Zhu et al. (2024).

A parallel survey strand focuses on how to unify KGs with LLMs. Pan et al. (2024) frames the space as (i) KG-enhanced LLMs, (ii) LLM-augmented KGs, and (iii) synergized LLM+KG systems, and surveys representative techniques across these directions. Surveys of retrieval-augmented generation systematize design choices for retrieval, augmentation, and answer synthesis in LLM pipelines Gao et al. (2023), while GraphRAG surveys focus specifically on graph-structured indexing and retrieval that leverages entity and relation structure for more precise context construction Peng et al. (2024). Finally, surveys of LLM-based multi-agent systems analyze how collections of LLM-driven agents coordinate roles, communicate, and use tools and environments, and they summarize emerging benchmarks for evaluating such systems Guo et al. (2024).

KG4ESG draws on these methodological insights but instantiates them in an ESG setting where cross-framework semantics, qualification (time/unit/scope/method), provenance, and auditability are first-class constraints.

Appendix B. Related Work Beyond Prior Surveys

This section complements Appendix A by highlighting primary technical threads that are most closely adjacent to KG4ESG's end-to-end framing. The emphasis is on concrete systems and artifacts

 doi:10.20944/preprints202602.1970.v2

(rather than survey/meta-review work) that instantiate parts of the **Data**→**KG** and **KG**→**App** pipeline under standards, qualification, provenance, and auditability constraints.

Appendix B.1. Disclosure Structuring & Cross-Framework Indicator Alignment

A growing set of ESG systems uses KGs as intermediate representations for turning narrative disclosures into comparable, framework-grounded indicators. These pipelines typically perform schema-conditioned structuring and alignment while retaining fine-grained evidence pointers (clauses/tables/spans) to support reviewer-facing traceability and re-auditing as standards evolve Yu et al. (2025); Bronzini et al. (2024); Zhou et al. (2024); Usmanova and Usbeck (2024). Closely related work emphasizes indicator semantics and computation provenance (units, boundaries, and method metadata) so that downstream analytics remain computable rather than merely retrievable Diamantini et al. (2022); Markovic et al. (2023); Diamantini et al. (2025). Recent expert-guided, LLM-empowered platforms for sustainability standards knowledge graphs further reinforce the importance of standards-centric graph infrastructures at this layer He et al. (2026a). KG4ESG differs by organizing these contributions as an atlas across ESG-RFM foci and by explicitly coupling construction choices to the behavior of downstream language interfaces.

Appendix B.2. Compliance- & Governance-Oriented KGs

In regulatory and governance deployments, the central representational problem is often clause-level modeling of obligations, scope, exceptions, and cross-references, together with provenance structures that keep the evidence chain inspectable. Representative work includes compliance-oriented legal/policy graphs Ershov (2023); Chung et al. (2025), audit-facing LLM+KG systems that answer with evidence-backed traces Tan et al. (2025); Jomraj et al. (2025), and privacy/governance graphs that prioritize verifiable retrieval and justification under policy constraints Garza et al. (2024); Hernandez et al. (2025); Tauqeer et al. (2022). These lines motivate KG4ESG's focus on provenance and qualification as first-class constraints, because they define the interaction contract between extraction, reasoning, and external review.

Appendix B.3. Graph-Grounded LLM Interfaces & Validation-First Workflows

Recent systems increasingly treat graph-grounded interfaces (KGQA, text-to-query, GraphRAG-style context construction) as the default access pattern for narrative-heavy ESG evidence, but pair them with explicit validation and conflict reporting to prevent silent semantic drift. In practice, this means typed outputs with constraint checks (e.g., units, scope, temporal validity), provenance enforcement, and replayable traces when tools or agents are used Yu et al. (2025); Cai et al. (2025); Jomraj et al. (2025); Ushio et al. (2025). At the application layer, operating-system-style proposals for planetary climate adaptation likewise point toward more integrated end-to-end application environments He et al. (2026b). KG4ESG abstracts these trends into P3 (LLM-assisted structuring/alignment) and P4 (agentic/tool-using workflows) so that reporting and evaluation practices become comparable across ESG topics and application contexts.

Appendix B.4. Monitoring, Controversy Intelligence, & Claim Verification

For monitoring and risk intelligence, KGs often act as event- and entity-centric indices over noisy narrative signals such as news, web pages, and reports. Works in this direction construct or align graphs for controversy tracking and ESG risk intelligence Angioni et al. (2024); Hassan Nassar et al. (2025); Iwata et al. (2025), and build claim- or fact-centric structures for transparent evidence-backed verification DeBellis et al. (2025); Kaoukis et al. (2025); Burel and Alani (2025). Compared to disclosure structuring, these settings face more severe uncertainty and source conflict, further amplifying the need for explicit provenance, versioning, and "show the evidence" interaction contracts—themes that recur in KG4ESG's meta-analysis and future research agenda.

 doi:10.20944/preprints202602.1970.v2

Appendix C. Construction of the ESG Research Focus Map (ESG-RFM)

Figure A1 provides a visual overview of the three-step workflow used to construct and harmonize the ESG-RFM.

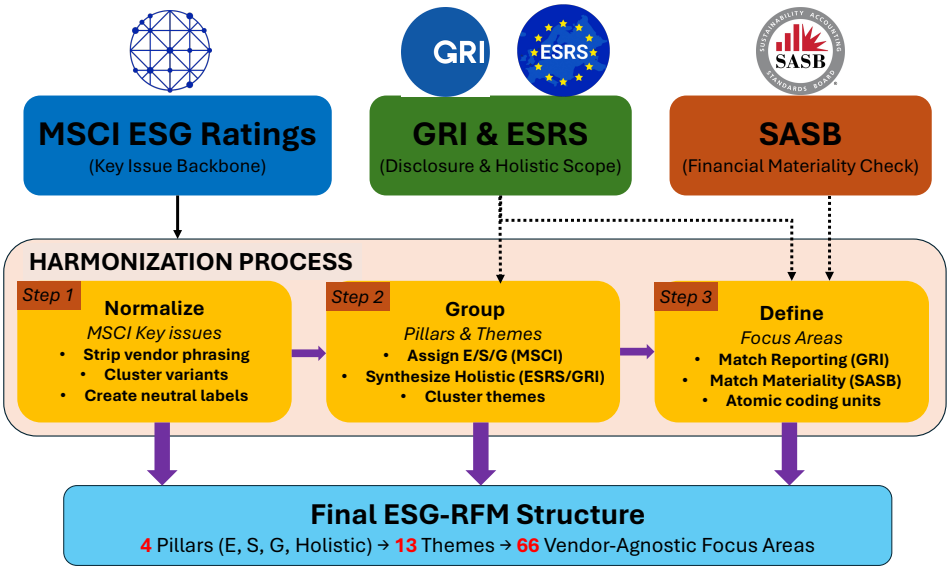

Figure A1. Visual overview of the ESG-RFM construction process. The process involves three main steps: (1) Normalizing MSCI Key Issues, (2) Grouping them into pillars and themes with holistic input from GRI and ESRS, and (3) Defining focus areas by matching with GRI reporting and SASB materiality. This results in the final ESG-RFM structure comprising 4 pillars, 13 themes, and 66 focus areas.

Appendix C.1. Source Frameworks and Priority Order

We integrate one widely used ESG rating methodology (MSCI[1]) with three prominent ESG standards families (GRI[2], ESRS[3], SASB[4]) to construct the ESG-RFM. These sources differ in purpose and scope, and this shapes the hierarchy we apply in organizing topics.

MSCI as the structural backbone MSCI's ESG Ratings assess company performance on environmental, social, and governance issues relative to peers, focusing on how exposed a company is to ESG risks and how effectively it manages those risks. The methodology centers on a hierarchy of ESG topics (pillar → theme → key issue), and is broadly adopted by investors for comparative assessment of corporate sustainability performance. We use MSCI's **Environmental, Social, and Governance** pillar assignments as the structural backbone of the map and as the starting point for topic normalization and cross-framework alignment.

GRI for detailed vocabulary and reporting practice The Global Reporting Initiative (GRI) Standards provide comprehensive guidance for sustainability reporting on a wide range of environmental, social, and governance topics. GRI's standards support organizations in identifying material sustainability issues and disclosing performance using standardized indicators and terminology, which enhances clarity and comparability in reporting. Because GRI elaborates specific indicators and reporting expectations, we use its standards to refine and articulate the terminology of broader ESG topics, anchoring them in widely recognized disclosure practice.

ESRS for double materiality and holistic expansion The European Sustainability Reporting Standards (ESRS), developed under the Corporate Sustainability Reporting Directive (CSRD), require companies to assess sustainability matters through a double materiality lens, considering both how sustainability issues affect the company and how the company's activities impact society and the environment.

[1] https://www.msci.com/data-and-analytics/sustainability-solutions/esg-ratings
[2] https://www.globalreporting.org/standards/
[3] https://www.efrag.org/en/sustainability-reporting
[4] https://www.ifrs.org/issued-standards/sasb-standards/

 doi:10.20944/preprints202602.1970.v2

This dual approach goes beyond a solely investor-centric perspective by integrating both financial and impact materiality. ESRS's structured topical standards allow broad issues to be conceptually unpacked into distinct material components and support formulation of broader themes that extend beyond the core three pillars of ESG.

SASB as a financial-materiality lens The SASB Standards (now maintained by the IFRS Foundation) provide industry-based disclosure standards that focus on sustainability issues most likely to have material impacts on a company's financial performance. Rather than adopting SASB's detailed industry-level standards wholesale, we use SASB's emphasis on financial materiality as a consistency check to ensure that theme groupings in the map align with considerations that matter for investor decision-making.

Across all four pillars of ESG-RFM, the integration of these sources ensures broad coverage: MSCI defines the structural backbone of environmental, social, and governance categories; GRI supplies reporting terminology and disclosure scope; ESRS embeds double materiality and supports extension into expansive sustainability concepts; and SASB reinforces financial-materiality interpretation for thematic coherence. This priority structure is reflected in the crosswalk Table A1 and Table A2, which map ESG-RFM labels to their upstream sources.

Appendix C.2. ESG-RFM Construction & Harmonization Process

The harmonization proceeds in three steps.

Step 1: Normalizing MSCI key issues into neutral labels

We begin by taking the full set of MSCI ESG Ratings key issues as the initial universe of issuer-level ESG topics and normalizing them into a **vendor-agnostic pool of candidate labels**. This normalization removes framework-specific phrasing and consolidates semantically equivalent issues into neutral ESG focus candidates, while preserving the original **Environmental, Social, and Governance** pillar assignments defined by MSCI. The result is a standardized topic layer that retains MSCI's coverage and structural backbone, while enabling consistent cross-framework alignment and subsequent harmonization with ESRS, GRI, SASB, and the broader ESG literature.

Step 2: Grouping topics into pillars and themes

Next, we organize the normalized candidate topics into a **four-pillar structure**, with a set of **themes** defined within each pillar. Topics originating from **MSCI Environmental, Social, and Governance key issues** retain their original pillar assignments and are mapped directly to the **Environmental**, **Social**, and **Governance** pillars, respectively. In addition, we introduce a **Holistic** pillar to capture *cross-cutting concepts* that are not modeled as a separate pillar in MSCI but are explicitly articulated in **ESRS**, **GRI**, **SASB**, and the broader ESG literature. This pillar covers concepts such as *materiality analysis*, *just transition*, *circular-economy perspectives*, and *systemic sustainability models*.

Within each pillar, we define **themes** by clustering candidate topics that are *conceptually related* and that form *coherent blocks* in disclosure practice and rating methodologies across MSCI, GRI, ESRS, and SASB. While MSCI also defines themes, we apply vendor-agnostic naming and (where needed) slight re-scoping to support cross-standard alignment and to accommodate the additional **Holistic** pillar.

The **Environmental** pillar is organized into four themes: (i) *Climate Change*, (ii) *Natural Capital*, (iii) *Pollution, Waste and Circularity*, and (iv) *Sustainability Solutions and Technologies*.

The **Social** pillar comprises four themes: (i) *Human Capital*, (ii) *Product Responsibility and Customer Safety*, (iii) *Community and Rights Risks*, and (iv) *Inclusive Solutions and Social Access*.

The **Governance** pillar is structured around two themes: (i) *Corporate Governance* and (ii) *Corporate Conduct and Integrity*.

Finally, the **Holistic** pillar is divided into three themes: (i) *Integrated ESG Assessment*, (ii) *Sustainable Development Frameworks*, and (iii) *Systemic Sustainability Models*. These themes draw primarily on **ESRS** and **GRI**, with additional support from **SASB** and recurring practices observed in the academic literature, rather than from standalone MSCI key issues.

 doi:10.20944/preprints202602.1970.v2

At this stage, we use **GRI** and **ESRS** primarily to *check that theme boundaries correspond to coherent disclosure blocks*. For instance, MSCI's "Water Stress" key issue is grouped with other water- and ecosystem-related topics under *Natural Capital*, in line with **GRI 303** (Water and Effluents) and **ESRS E3** (Water and marine resources). The resulting mapping between KG4ESG pillars and themes and representative MSCI, GRI, ESRS, and SASB topics is summarized in Table A1.

Step 3: Defining focus areas as units of analysis

Within each theme, we define **focus** as the *atomic units of analysis* used for querying and coding the KG4ESG corpus. Each focus area is anchored in one or more **MSCI key issues**, or in cross-cutting MSCI metrics for the **Environmental**, **Social**, and **Governance** pillars, which serve as the structural backbone of the taxonomy.

For the **Holistic** pillar, focus areas are instead grounded in *explicit cross-cutting constructs* defined in **ESRS** and **GRI**, with additional support from **SASB** and recurring patterns observed in the academic literature, rather than being tied to a single MSCI key issue.

Across all pillars, each focus area is required to have at least one clear correspondence in **GRI** or **ESRS**, ensuring well-defined disclosure scope and impact materiality, and typically aligns with one or more **SASB** general issue categories where applicable, providing a complementary perspective on financial materiality.

Table A2 presents each ESG-RFM focus as one row, with the right-hand columns showing the closest upstream topics in MSCI, GRI, ESRS and SASB. This mapping is deliberately compact and non-exhaustive: it records the *main anchors* that informed the KG4ESG label rather than attempting to list all possible standard references.

Table A1. Conceptual crosswalk between the ESG-RFM (pillar–theme; abbreviations per Table A3) and representative topics in MSCI, GRI, ESRS and SASB. Multiple upstream issues can map to a single KG4ESG theme and vice versa. Wording in the right-hand columns is summarised and does not reproduce the original frameworks.

Pillar	Theme	MSCI key issues (representative)	GRI / ESRS topics (representative)	SASB issue categories (representative)
E	`Clim Chg`	Carbon Emissions; Climate Change Vulnerability; Financing Environmental Impact; Product Carbon Footprint	GRI: GRI 102: Climate Change 2025; GRI 103: Energy 2025 (legacy: GRI 302/305, GRI 201-2); ESRS: E1 Climate change (transition and physical risks, decarbonisation)	Environment: GHG Emissions; Energy Management; Physical impacts of climate change
E	`Nat Cap`	Biodiversity & Land Use; Raw Material Sourcing; Water Stress; Materials Use & Resource Efficiency (via resource-related indicators)	GRI: Materials, Water and Effluents, Biodiversity, Supplier Environmental Assessment; ESRS: E3 Water and marine resources, E4 Biodiversity and ecosystems, E5 Resource use and circular economy, S2 Workers in the value chain	Environment: Water & Wastewater Management; Ecological Impacts; Resource Efficiency; Supply Chain Management
E	`Pol Was Ciru`	Toxic Emissions & Waste; Packaging Material & Waste; Electronic Waste; cross-cutting compliance indicators	GRI: Emissions, Waste, Environmental Compliance; ESRS: E2 Pollution, E5 Resource use and circular economy	Environment: Air Quality; Waste & Hazardous Materials Management; Product Design & Lifecycle Management
E	`Sust Sol Tech`	Opportunities in Clean Tech; Opportunities in Renewable Energy; Opportunities in Green Building	GRI: Energy, Emissions, Indirect Economic Impacts; ESRS: E1 Climate change, E5 Resource use and circular economy	Environment: GHG Emissions; Energy Management; Business Model & Innovation (low-carbon products)
S	`Hum Cap`	Health & Safety; Human Capital Development; Labor Management; Supply Chain Labor Standards (plus diversity- and job-quality-related metrics)	GRI: Employment, Labour/Management Relations, Occupational Health and Safety, Training and Education, Diversity and Equal Opportunity, Non-discrimination, Market Presence, Supplier Social Assessment; ESRS: S1 Own workforce, S2 Workers in the value chain	Human Capital: Employee Health & Safety; Labor Practices; Employee Engagement, Diversity & Inclusion; Supply Chain Management
S	`Prod Resp Cust Safe`	Chemical Safety; Consumer Financial Protection; Privacy & Data Security; Product Safety & Quality; Health & Demographic Risk; Responsible Marketing & Product Labeling; Responsible Investment	GRI: Product Quality and Safety, Marketing and Labeling, Customer Health and Safety, Customer Privacy, Financial sector supplements; ESRS: S4 Consumers and end-users, E2 Pollution (hazardous substances)	Social Capital: Product Quality & Safety; Customer Welfare; Customer Privacy; Access & Affordability; Leadership & Governance (Financials)

 doi:10.20944/preprints202602.1970.v2

Table A1. *Cont.*

Pillar	Theme	MSCI key issues (representative)	GRI / ESRS topics (representative)	SASB issue categories (representative)
S	Com Rig Risks	Community Relations; Controversial Sourcing; human-rights-related factors across social key issues	GRI: Local Communities, Human Rights series, Supplier Social Assessment; ESRS: S2 Workers in the value chain, S3 Affected communities	Social Capital: Human Rights & Community Relations; Supply Chain Management
S	Inc Sol Soc Acc	Access to Finance; Access to Health Care; Access to Communications; Opportunities in Nutrition & Health	GRI: Product and Service Labelling, Indirect Economic Impacts, sector standards (financial, health, telecom); ESRS: S3 Affected communities, S4 Consumers and end-users	Social Capital: Access & Affordability; Product Quality & Safety; Business Model & Innovation (inclusive products)
G	Corp Gov	Board; Pay; Ownership & Control; Accounting	GRI: Governance structure and composition, Remuneration; ESRS: ESRS 2 General disclosures (governance)	Leadership & Governance: Board structure & oversight; Incentive structure; Systemic risk management
G	Corp Cond Int	Business Ethics; Tax Transparency; Public Policy; cross-cutting compliance and controversy indicators	GRI: Anti-corruption, Anti-competitive Behaviour, Tax, Socioeconomic Compliance, Procurement Practices, Customer Privacy; ESRS: G1 Business conduct, ESRS 2 Governance disclosures	Leadership & Governance: Business Ethics; Competitive Behaviour; Tax Transparency; Regulatory environment; Social Capital: Customer Privacy; Supply Chain Management
H	ESG Int Ass	Overall ESG performance and rating; ESG risk management; data integration; materiality assessment; scenario analysis; reporting & disclosure; controversies	GRI: Material topics and general disclosures; ESRS: ESRS 1 General requirements, ESRS 2 General disclosures and cross-topic matters	Cross-cutting: Topic selection and materiality; Risk Management; Data security & reporting across all SASB dimensions
H	SD Frame	SDG alignment; impact investing; reporting frameworks; ESG taxonomies and standards alignment	GRI: SDG mapping and sector-specific standards; ESRS: references to EU Taxonomy and SDGs, interoperability guidance	Cross-cutting: SDG alignment; Impact metrics; Framework mapping (e.g. GRI / SASB / TCFD)
H	Sys Sust Mod	Circular economy; holistic supply chain sustainability; just transition and socio-ecological transformation	GRI: Materials, Waste, Economic, Social topics; ESRS: E5 Resource use and circular economy, S1–S4 social standards	Environment: Resource Efficiency & Waste; Social Capital: Community Relations; Human Capital: Labor Practices; Business Model & Innovation

Table A2. Focus-level crosswalk between the ESG-RFM (pillar–theme–focus) and representative topics in MSCI, GRI, ESRS and SASB. Multiple upstream issues can map to a single KG4ESG focus area and vice versa. Wording in the right-hand columns is summarized and does not reproduce the original frameworks.

Pillar	Theme	Focus	MSCI topics (representative)	GRI / ESRS topics (representative)	SASB issue categories (representative)
E	Clim Chg	GHG	Carbon Emissions; (partly) Product Carbon Footprint	GRI 102: Climate Change 2025; GRI 103: Energy 2025 (legacy: GRI 302: Energy; GRI 305: Emissions); ESRS E1 Climate change (GHG emissions, energy use)	Environment: GHG Emissions; Energy Management
E	Clim Chg	Phys Clim Risk	Climate Change Vulnerability	GRI 201: Economic Performance (climate-related financial implications; legacy: GRI 201-2); GRI 102: Climate Change 2025; ESRS E1 Climate change (physical risks, scenario analysis)	Environment: Physical impacts of climate change; cross-cutting climate risk metrics
E	Clim Chg	Clim Fin	Financing Environmental Impact	GRI 201: Economic Performance; GRI 203: Indirect Economic Impacts; GRI 102: Climate Change 2025; ESRS E1 Climate change (transition plans, financed emissions); ESRS E5 Resource use and circular economy (green finance for circular solutions)	Environment & Business Model/Innovation: GHG Emissions and Energy Management for financed activities; Product Design & Lifecycle for green products
E	Clim Chg	Energy Mix	Primarily captured within Carbon Emissions key issue (energy intensity and fuel mix metrics)	GRI 103: Energy 2025; GRI 102: Climate Change 2025 (legacy: GRI 302/305); ESRS E1 Climate change (energy consumption and mix)	Environment: Energy Management; Fuel Management
E	Clim Chg	Prod Foot	Product Carbon Footprint	GRI 102: Climate Change 2025 (value-chain emissions); GRI 301: Materials (legacy: GRI 305: Emissions); ESRS E1 Climate change (value-chain emissions)	Environment & Business Model/Innovation: GHG Emissions; Product Design & Lifecycle Management
E	Nat Cap	Biodiv LU	Biodiversity & Land Use	GRI 304: Biodiversity; ESRS E4 Biodiversity and ecosystems	Environment: Ecological Impacts
E	Nat Cap	Raw Src	Raw Material Sourcing	GRI 301: Materials; GRI 204: Procurement Practices; GRI 308: Supplier Environmental Assessment; ESRS E5 Resource use and circular economy; ESRS S2 Workers in the value chain	Environment & Social Capital: Ecological Impacts; Supply Chain Management
E	Nat Cap	Res Use Mat Eff	Raw Material Sourcing; (indirectly) Electronic Waste and Packaging Material & Waste where resource efficiency affects downstream waste	GRI 301: Materials; GRI 302: Energy; GRI 306: Waste; ESRS E5 Resource use and circular economy	Environment & Business Model/Innovation: Waste & Hazardous Materials Management; Product Design & Lifecycle Management

 doi:10.20944/preprints202602.1970.v2

Table A2. *Cont.*

Pillar	Theme	Focus	MSCI topics (representative)	GRI / ESRS topics (representative)	SASB issue categories (representative)
E	Nat Cap	Supp Env DD	No dedicated key issue; mainly reflected through Raw Material Sourcing, Electronic Waste and sector-specific supply-chain indicators across the Environmental pillar	GRI 308: Supplier Environmental Assessment; GRI 204: Procurement Practices; ESRS E2–E5 environmental standards and ESRS S2 Workers in the value chain	Environment & Social Capital: Supply Chain Management; Ecological Impacts
E	Nat Cap	Water SS	Water Stress	GRI 303: Water and Effluents; ESRS E3 Water and marine resources	Environment: Water & Wastewater Management
E	Pol Was Ciru	EWaste	Electronic Waste	GRI 301: Materials; GRI 306: Waste; GRI 303: Water and Effluents (where relevant); ESRS E2 Pollution; ESRS E5 Resource use and circular economy	Environment: Waste & Hazardous Materials Management; Product Design & Lifecycle Management
E	Pol Was Ciru	Pack Waste	Packaging Material & Waste	GRI 301: Materials; GRI 306: Waste; ESRS E2 Pollution; ESRS E5 Resource use and circular economy	Environment: Waste & Hazardous Materials Management; Product Design & Lifecycle Management
E	Pol Was Ciru	Env Comp	No single dedicated key issue; environmental compliance and management systems are cross-cutting metrics used to assess performance on all Environmental key issues	GRI 307: Environmental Compliance; GRI 305: Emissions; GRI 306: Waste; ESRS 2 General disclosures (compliance with laws and regulations); ESRS E2 Pollution	Environment & Leadership/Governance: Air Quality; Waste & Hazardous Materials Management; Systemic risk and compliance management
E	Pol Was Ciru	Haz Waste	Toxic Emissions & Waste	GRI 305: Emissions; GRI 306: Waste; GRI 303: Water and Effluents; GRI 307: Environmental Compliance; ESRS E2 Pollution	Environment: Air Quality; Waste & Hazardous Materials Management
E	Sust Sol Tech	Clean Tech	Opportunities in Clean Tech	GRI 302: Energy; GRI 305: Emissions; GRI 201: Economic Performance (green revenues); ESRS E1 Climate change (transition plans, low-carbon products); ESRS E5 Resource use and circular economy	Environment & Business Model/Innovation: GHG Emissions; Energy Management; Product Design & Lifecycle Management (low-carbon solutions)
E	Sust Sol Tech	Green Bldg	Opportunities in Green Building	GRI 302: Energy; GRI 305: Emissions; GRI 303: Water and Effluents; GRI 306: Waste; ESRS E1 Climate change (building energy performance); ESRS E3 Water and marine resources; ESRS E5 Resource use and circular economy	Environment: Energy Management; Water & Wastewater Management; Waste & Hazardous Materials Management; Business Model/Innovation: Product Design & Lifecycle Management (building design)
E	Sust Sol Tech	Re Eneg	Opportunities in Renewable Energy	GRI 302: Energy; GRI 305: Emissions; GRI 201: Economic Performance (renewable investments); ESRS E1 Climate change (renewable energy generation and use)	Environment: GHG Emissions; Energy Management; Business Model/Innovation: Product Design & Lifecycle Management (renewable offerings)
S	Hum Cap	OHS	Health & Safety	GRI 403: Occupational Health and Safety; GRI 401: Employment; ESRS S1 Own workforce (health and safety); ESRS S2 Workers in the value chain (OHS in supply chains)	Human Capital: Employee Health & Safety
S	Hum Cap	Learn Dev	Human Capital Development	GRI 404: Training and Education; GRI 401: Employment; ESRS S1 Own workforce (skills, training, career development)	Human Capital: Labor Practices; Employee Engagement, Diversity & Inclusion (training, retention)
S	Hum Cap	Work Man Rel	Labor Management	GRI 402: Labour/Management Relations; GRI 401: Employment; ESRS S1 Own workforce; ESRS S2 Workers in the value chain	Human Capital: Labor Practices; Employee Engagement, Diversity & Inclusion
S	Hum Cap	Job Qual	No dedicated key issue; mainly reflected in Health & Safety and Labor Management key issues and related Human Capital metrics	GRI 401: Employment; GRI 403: Occupational Health and Safety; GRI 404: Training and Education; ESRS S1 Own workforce (fair wages, work conditions)	Human Capital: Labor Practices; Employee Health & Safety; Employee Engagement, Diversity & Inclusion
S	Hum Cap	DEI	Assessed as part of Human Capital theme (e.g. workforce diversity metrics under Labor Management and Human Capital Development)	GRI 405: Diversity and Equal Opportunity; GRI 406: Non-discrimination; ESRS S1 Own workforce (diversity, equal treatment)	Human Capital: Employee Engagement, Diversity & Inclusion
S	Hum Cap	FoA Coll B	Assessed primarily within Labor Management key issue (unionization, collective bargaining) and related controversy indicators	GRI 407: Freedom of Association and Collective Bargaining; GRI 402: Labour/Management Relations; ESRS S1 Own workforce and ESRS S2 Workers in the value chain (freedom of association and CB)	Human Capital & Social Capital: Labor Practices; Supply Chain Management
S	Hum Cap	Loc Hir Mar Pre	Addressed through market-presence and employment-related metrics within the Human Capital theme, rather than a standalone key issue	GRI 202: Market Presence; GRI 203: Indirect Economic Impacts; GRI 401: Employment; ESRS S1 Own workforce; ESRS S3 Affected communities	Human Capital & Social Capital: Labor Practices; Human Rights & Community Relations
S	Hum Cap	Lab Stan SC	Supply Chain Labor Standards	GRI 414: Supplier Social Assessment; GRI 408: Child Labor; GRI 409: Forced or Compulsory Labour; ESRS S2 Workers in the value chain; ESRS S3 Affected communities	Social Capital & Human Capital: Supply Chain Management; Human Rights & Community Relations
S	Prod Resp Cust Safe	Chem Safe	Chemical Safety	GRI 416: Customer Health and Safety; GRI 306: Waste; GRI 303: Water and Effluents (discharges of hazardous substances); ESRS E2 Pollution; ESRS S4 Consumers and end-users	Environment & Social Capital: Waste & Hazardous Materials Management; Product Quality & Safety; Customer Welfare

 doi:10.20944/preprints202602.1970.v2

Table A2. *Cont.*

Pillar	Theme	Focus	MSCI topics (representative)	GRI / ESRS topics (representative)	SASB issue categories (representative)
S	Prod Resp Cust Safe	Fin Prot	Consumer Financial Protection	GRI 416: Customer Health and Safety; GRI 417: Marketing and Labeling; GRI 418: Customer Privacy; financial sector GRI supplements; ESRS S4 Consumers and end-users (fair treatment, product suitability)	Social Capital: Customer Welfare; Product Quality & Safety; Customer Privacy & Data Security
S	Prod Resp Cust Safe	Data Sec	Privacy & Data Security	GRI 418: Customer Privacy; GRI 419: Socioeconomic Compliance (for data-related fines); ESRS S4 Consumers and end-users (privacy, security); ESRS G1 Business conduct (data and digital ethics)	Social Capital & Leadership/Governance: Customer Privacy; Data Security; Business Ethics
S	Prod Resp Cust Safe	Prod Qual	Product Safety & Quality	GRI 416: Customer Health and Safety; GRI 417: Marketing and Labelling; ESRS S4 Consumers and end-users (product safety and quality)	Social Capital: Product Quality & Safety; Customer Welfare; Selling Practices & Product Labeling
S	Prod Resp Cust Safe	Heal Vuln	Health & Demographic Risk	GRI 416: Customer Health and Safety; GRI 203: Indirect Economic Impacts; GRI 413: Local Communities; ESRS S3 Affected communities; ESRS S4 Consumers and end-users	Social Capital: Customer Welfare; Human Rights & Community Relations
S	Prod Resp Cust Safe	Resp Mkt	Responsible Marketing & Product Labeling	GRI 417: Marketing and Labeling; GRI 416: Customer Health and Safety; ESRS S4 Consumers and end-users (marketing, communication, fair treatment)	Social Capital: Selling Practices & Product Labeling; Customer Welfare
S	Prod Resp Cust Safe	RSInv	Responsible Investment	GRI 201: Economic Performance; GRI 203: Indirect Economic Impacts; financial-sector GRI standards; ESRS E1–E5 and S1–S4 as underlying sustainability impacts considered in investment products	Cross-cutting across SASB dimensions for financials (e.g. GHG Emissions, Customer Welfare, Human Capital, Business Ethics)
S	Com Rig Risks	Com Rel	Community Relations	GRI 413: Local Communities; GRI 411: Rights of Indigenous Peoples; ESRS S3 Affected communities	Social Capital: Human Rights & Community Relations
S	Com Rig Risks	HRDD	Addressed across Social key issues (e.g. Labor Management, Supply Chain Labor Standards, Community Relations) rather than a single named key issue	GRI 408: Child Labor; GRI 409: Forced or Compulsory Labor; GRI 410: Security Practices; GRI 414: Supplier Social Assessment; ESRS S1–S3 social standards; ESRS G1 Business conduct	Social Capital & Human Capital: Human Rights & Community Relations; Labor Practices; Supply Chain Management
S	Com Rig Risks	Risk Src	Controversial Sourcing	GRI 204: Procurement Practices; GRI 308: Supplier Environmental Assessment; GRI 414: Supplier Social Assessment; ESRS S2 Workers in the value chain; ESRS S3 Affected communities; ESRS E3/E5 (high-risk raw materials)	Environment & Social Capital: Supply Chain Management; Ecological Impacts; Human Rights & Community Relations
S	Inc Sol Soc Acc	Fin Acc	Access to Finance	GRI 203: Indirect Economic Impacts; GRI 416: Customer Health and Safety; sector-specific financial services Standards; ESRS S4 Consumers and end-users (access and affordability of financial services)	Social Capital: Access & Affordability; Customer Welfare
S	Inc Sol Soc Acc	Heal Acc	Access to Health Care	GRI 416: Customer Health and Safety; GRI 203: Indirect Economic Impacts; GRI 413: Local Communities; ESRS S3 Affected communities; ESRS S4 Consumers and end-users	Social Capital: Access & Affordability; Customer Welfare; Product Quality & Safety
S	Inc Sol Soc Acc	Dig Incl	Access to Communications	GRI 203: Indirect Economic Impacts; GRI 413: Local Communities; GRI 416: Customer Health and Safety; ESRS S1 Own workforce (digital inclusion of employees); ESRS S4 Consumers and end-users (digital access, inclusion)	Social Capital: Access & Affordability; Customer Welfare; Data Security
S	Inc Sol Soc Acc	Nutri Heal	Opportunities in Nutrition & Health	GRI 416: Customer Health and Safety; GRI 203: Indirect Economic Impacts; GRI 413: Local Communities; ESRS S3 Affected communities; ESRS S4 Consumers and end-users	Social Capital: Customer Welfare; Access & Affordability; Product Quality & Safety
G	Corp Gov	Board Ovr	Board	GRI 2: General Disclosures (governance structure and composition); ESRS 2 General disclosures (governance, strategy and management of impacts and risks)	Leadership & Governance: Board structure and oversight; Systemic risk management
G	Corp Gov	Exec Pay	Pay	GRI 2: General Disclosures (remuneration policies and practices); GRI 201: Economic Performance (links to variable remuneration); ESRS 2 General disclosures (incentives linked to sustainability performance)	Leadership & Governance: Incentive and remuneration alignment; Business Ethics
G	Corp Gov	Own Ctrl	Ownership & Control	GRI 2: General Disclosures (ownership and control); GRI 207: Tax (transparency on structures); ESRS 2 General disclosures; ESRS G1 Business conduct (transparency, beneficial ownership)	Leadership & Governance: Business Ethics; Competitive Behavior; Systemic risk management
G	Corp Gov	Fin RepQ	Accounting	GRI 201: Economic Performance; GRI 2: General Disclosures (internal controls, audit, governance); ESRS 1 General requirements and ESRS 2 General disclosures (link between sustainability and financial information)	Leadership & Governance: Systemic risk management; Business Ethics

 doi:10.20944/preprints202602.1970.v2

Table A2. *Cont.*

Pillar	Theme	Focus	MSCI topics (representative)	GRI / ESRS topics (representative)	SASB issue categories (representative)
G	Corp Cond Int	Ethics AC	Business Ethics	GRI 205: Anti-corruption; GRI 206: Anti-competitive Behavior; GRI 419: Socioeconomic Compliance; ESRS G1 Business conduct	Leadership & Governance: Business Ethics; Competitive Behavior; Systemic risk management
G	Corp Cond Int	Tax Trans	Tax Transparency	GRI 207: Tax; GRI 201: Economic Performance; GRI 419: Socioeconomic Compliance; ESRS G1 Business conduct (tax governance, country-by-country reporting)	Leadership & Governance: Business Ethics (tax planning and transparency); Systemic risk management
G	Corp Cond Int	Pol Engage	Public Policy	GRI 415: Public Policy; GRI 201: Economic Performance (public subsidies); GRI 419: Socioeconomic Compliance; ESRS G1 Business conduct; ESRS 2 General disclosures (lobbying, political engagement)	Leadership & Governance: Business Ethics; Systemic risk management; Social Capital: Human Rights & Community Relations (policy impacts)
G	Corp Cond Int	Reg Comp	Reflected through cross-cutting compliance indicators rather than a named key issue (e.g. fines and controversies across Environmental, Social and Governance topics)	GRI 419: Socioeconomic Compliance; GRI 307: Environmental Compliance; GRI 2-27: Compliance with laws and regulations; ESRS 2 General disclosures (compliance and litigation)	Leadership & Governance: Business Ethics; Systemic risk and compliance management
G	Corp Cond Int	Resp Proc	Assessed mainly through Community Relations, Raw Material Sourcing and Supply Chain Labor Standards key issues, depending on impact type	GRI 204: Procurement Practices; GRI 308: Supplier Environmental Assessment; GRI 414: Supplier Social Assessment; GRI 413: Local Communities; ESRS S2 Workers in the value chain; ESRS E5 Resource use and circular economy	Environment, Human Capital & Social Capital: Supply Chain Management; Human Rights & Community Relations
G	Corp Cond Int	Fair Comp	Covered primarily within Business Ethics key issue (anti-competitive conduct, market abuse) and related controversy indicators	GRI 206: Anti-competitive Behavior; GRI 205: Anti-corruption; ESRS G1 Business conduct	Leadership & Governance: Competitive Behavior; Business Ethics
G	Corp Cond Int	Whistle	Included within Business Ethics key issue (whistleblower mechanisms, retaliation risks) and across governance controversy indicators	GRI 2: General Disclosures (ethics, integrity, speaking up); GRI 205: Anti-corruption; GRI 406: Non-discrimination; ESRS G1 Business conduct	Leadership & Governance: Business Ethics; Systemic risk and compliance management
G	Corp Cond Int	AI Ethics	Most closely linked to Privacy & Data Security and Business Ethics key issues, as well as to emerging "digital responsibility" assessments in ESG ratings	GRI 418: Customer Privacy; GRI 416: Customer Health and Safety (digital products); GRI 205: Anti-corruption; ESRS S4 Consumers and end-users (digital services, privacy and safety); ESRS G1 Business conduct (AI and digital conduct)	Social Capital & Leadership/Governance: Customer Privacy; Data Security; Business Ethics
H	ESG Int Ass	ESG Rate	Overall MSCI ESG rating derived from the full set of Environmental, Social and Governance key issues	GRI 1: Foundation; GRI 2: General Disclosures; GRI 3: Material Topics; full suite of topic Standards; ESRS 1 General requirements; ESRS 2 General disclosures	Cross-cutting across all Environment, Social Capital, Human Capital, Business Model & Innovation, and Leadership & Governance issues
H	ESG Int Ass	ESG Data	Data architecture and aggregation rules underlying the ESG Ratings methodology rather than a discrete key issue	GRI 2: General Disclosures (data quality, boundaries, restatements); GRI 3: Material Topics; ESRS 2 General disclosures (data, estimates, and value-chain coverage)	Leadership & Governance: Systemic risk management and disclosure quality; cross-cutting use of SASB indicators
H	ESG Int Ass	Econ Imp	Indirect economic and social effects captured across E, S and G key issues (e.g. community impacts, employment, innovation) rather than a separate key issue	GRI 201: Economic Performance; GRI 203: Indirect Economic Impacts; GRI 204: Procurement Practices; ESRS E1–E5 and S1–S4 (impacts, risks and opportunities on economy and society)	Cross-cutting: Business Model & Innovation (economic resilience); Social Capital and Human Capital (distributional and community effects)
H	ESG Int Ass	Dual Mat	Internal process for issue weighting and materiality mapping that determines the selection and weight of key issues in the overall rating	GRI 3: Material Topics (materiality determination); GRI 1: Foundation (impact materiality); ESRS 1 General requirements (double materiality) and ESRS 2 General disclosures (material impacts, risks and opportunities)	All SASB dimensions, via the SASB Materiality Map used to identify financially material issues
H	ESG Int Ass	ESG Scen	Climate and ESG risk assessment embedded in key issues such as Carbon Emissions, Climate Change Vulnerability, Business Ethics and others	GRI Climate Change and Energy Standards (GRI 102 Climate Change 2025 and revised GRI 302/305); GRI 201: Economic Performance (climate-related financial implications); ESRS E1 Climate change (risk and scenario analysis) and other topical ESRS standards	Environment & Leadership/Governance: Physical impacts of climate change; Systemic risk management; cross-cutting risk metrics
H	ESG Int Ass	ESG Rept	Underlying disclosure coverage and transparency assessments across all key issues and pillars	GRI 1, GRI 2 and GRI 3 (overall reporting architecture); all topic Standards; ESRS 1 General requirements; ESRS 2 General disclosures	Cross-cutting: all SASB issues to the extent disclosed using SASB metrics and narrative
H	ESG Int Ass	Green WH	Reflected mainly in Business Ethics and ESG controversy indicators (misleading claims, selective disclosure, controversies on sustainability claims)	GRI 2-22 and 2-27 (sustainable development strategy and compliance); GRI 417: Marketing and Labelling; GRI 419: Socioeconomic Compliance; ESRS 2 General disclosures; ESRS G1 Business conduct	Leadership & Governance: Business Ethics; Social Capital: Selling Practices & Product Labelling; Systemic risk and reputation management
H	ESG Int Ass	ESG Incid	MSCI ESG Controversies framework, which flags incidents across all Environmental, Social and Governance key issues	GRI 2-27: Compliance with laws and regulations; GRI 307: Environmental Compliance; GRI 419: Socioeconomic Compliance; ESRS 2 General disclosures (incidents, severe impacts) and topical ESRS incident-related requirements	Cross-cutting across all SASB issues, captured via incident, controversy and regulatory metrics

 doi:10.20944/preprints202602.1970.v2

Table A2. *Cont.*

Pillar	Theme	Focus	MSCI topics (representative)	GRI / ESRS topics (representative)	SASB issue categories (representative)
H	SD Frame	SDGs	MSCI SDG-alignment analytics linking key issues and revenues to SDG targets	GRI SDG mapping guidance (linking topic Standards to SDGs); ESRS references to EU Taxonomy and SDG alignment in ESRS 1 and ESRS 2	Cross-cutting: SASB SDG mapping linking general issue categories and industry metrics to SDG targets
H	SD Frame	Imp Inv	MSCI impact and thematic ESG indexes and solutions derived from key issue exposures (e.g. climate, natural capital, social themes)	GRI 201: Economic Performance; GRI 203: Indirect Economic Impacts; sector-specific financial services Standards; ESRS E1–E5 and S1–S4 as impact lenses for investment products	Cross-cutting: environment, social and governance issues that underpin impact investing taxonomies and metrics
H	SD Frame	Rept Std	Considered as upstream frameworks informing MSCI data inputs and methodologies rather than key issues per se	GRI 1, GRI 2, GRI 3 and all topic Standards (full GRI architecture); ESRS 1 and ESRS 2 as the EU sustainability reporting baseline	All SASB industry-specific Standards and general issue categories, used as reporting and disclosure benchmarks
H	SD Frame	Tax Map	Internal mapping of MSCI key issues to regulations and taxonomies (e.g. EU Taxonomy, SFDR), used in product construction	GRI taxonomy of topics; ESRS references to EU Taxonomy and other frameworks (in ESRS 1 and topical appendices)	Cross-cutting mapping between SASB issues, SDGs, TCFD and other sustainability frameworks
H	Sys Sust Mod	Circ Loop	Reflected mainly in Raw Material Sourcing, Electronic Waste, Packaging Material & Waste and related key issues on resource efficiency	GRI 301: Materials; GRI 306: Waste; GRI 303: Water and Effluents; GRI 305: Emissions; ESRS E5 Resource use and circular economy	Environment & Business Model/Innovation: Waste & Hazardous Materials Management; Product Design & Lifecycle Management; Resource efficiency
H	Sys Sust Mod	Just Trans	Cross-cutting across Carbon Emissions, Climate Change Vulnerability, Supply Chain Labor Standards and Community Relations key issues (transition impacts on workers and communities)	GRI 201: Economic Performance; GRI 203: Indirect Economic Impacts; GRI 413: Local Communities; GRI 408/409 on labour rights; ESRS E1 Climate change (just transition references); ESRS S1–S3 social standards	Environment, Human Capital & Social Capital: GHG Emissions; Labor Practices; Human Rights & Community Relations
H	Sys Sust Mod	SC Sust	Combination of Raw Material Sourcing, Supply Chain Labor Standards, Electronic Waste, Packaging Material & Waste, Controversial Sourcing and other value-chain key issues	GRI 204: Procurement Practices; GRI 308: Supplier Environmental Assessment; GRI 414: Supplier Social Assessment; GRI 303, 305 and 306 for environmental impacts; ESRS S2 Workers in the value chain; ESRS S3 Affected communities; ESRS E2–E5 environmental standards	Environment, Human Capital & Social Capital: Supply Chain Management; Ecological Impacts; Human Rights & Community Relations

Appendix C.3. Pillar–Theme Mapping Between the ESG-RFM & ESG Frameworks

To document the construction of the ESG-RFM, this subsection sets out how the pillar–theme structure aligns with upstream ESG standards and ratings. The KG4ESG labels are deliberately paraphrased and vendor-agnostic, so the mapping is indicative rather than strictly one-to-one. The right-hand columns of Table A1 summarise, in a compact and non-exhaustive way, the most closely related topics in MSCI, GRI, ESRS and SASB.

Appendix C.4. Focus-Level Mapping Between the ESG-RFM & ESG Frameworks

At the focus level, the ESG-RFM becomes more granular. In what follows, each focus is linked to the closest upstream topics across the four frameworks. MSCI ESG key issues remain the main starting point for issuer-level topics, but for many focus areas—especially social sub-topics and holistic concepts such as dual materiality or just transition—the most informative upstream anchors come from GRI, ESRS or SASB rather than from a single MSCI key issue. The long table below provides a compact, non-exhaustive crosswalk. Where no dedicated MSCI key issue exists, the MSCI column explicitly notes that the concept is covered via cross-cutting metrics or through several key issues combined.

 doi:10.20944/preprints202602.1970.v2

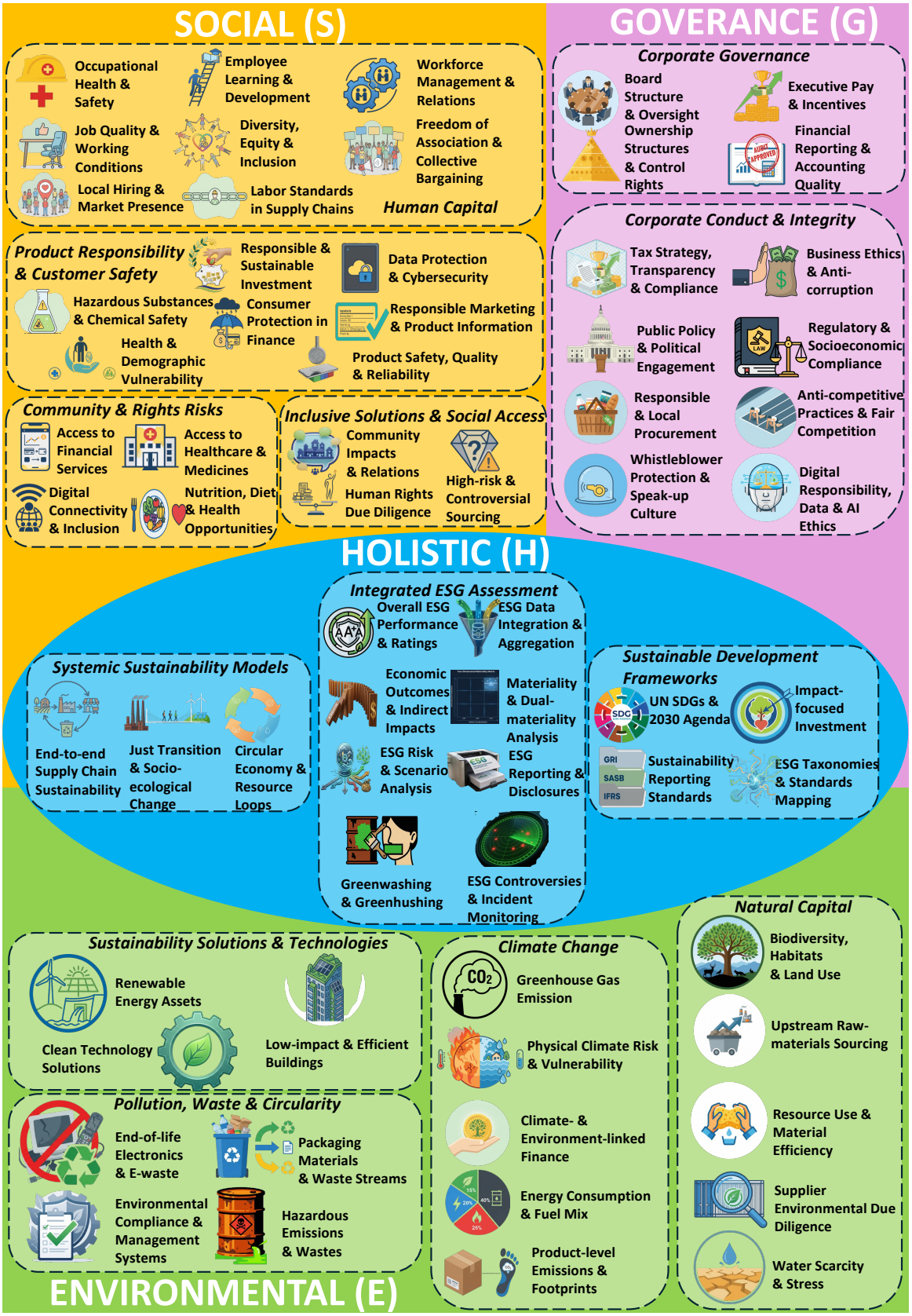

Figure A2. The ESG Research Focus Map (ESG-RFM). The map organizes ESG research into a three-level hierarchy of **pillars**, **themes**, and **research focus**, integrating ESG rating methodologies and reporting standards to provide a unified semantic structure for corpus annotation, cross-study comparison, and downstream knowledge graph construction.

 doi:10.20944/preprints202602.1970.v2

Appendix C.5. Abbreviation Dictionary for Pillars, Themes, and Focus

Table A3. Pillar and theme abbreviations (E/S/G/H).

Pillar (abbr.)	Theme	Theme abbr.
Environmental (E)	Climate Change	`ClimChg`
	Natural Capital	`NatCap`
	Pollution, Waste & Circularity	`PolWasCiru`
	Sustainability Solutions & Technologies	`SustSolTech`
Social (S)	Human Capital	`HumCap`
	Product Responsibility & Customer Safety	`ProdRespCustSafe`
	Community & Rights Risks	`ComRigRisks`
	Inclusive Solutions & Social Access	`IncSolSocAcc`
Governance (G)	Corporate Governance	`CorpGov`
	Corporate Conduct & Integrity	`CorpCondInt`
Holistic (H)	Integrated ESG Assessment	`ESGIntAss`
	Sustainable Development Frameworks	`SDFrame`
	Systemic Sustainability Models	`SysSustMod`

Table A4. Focus abbreviations keyed by `Pillar/Theme` (e.g., `E/ClimChg`, `S/HumCap`).

Pillar/Theme	Focus	Focus abbr.
`E/ClimChg`	Greenhouse Gas Emissions	`GHG`
`E/ClimChg`	Physical Climate Risk & Vulnerability	`PhysClimRisk`
`E/ClimChg`	Climate-& Environment-linked Finance	`ClimFin`
`E/ClimChg`	Energy Consumption & Fuel Mix	`EnergyMix`
`E/ClimChg`	Product-level Emissions & Footprints	`ProdFoot`
`E/NatCap`	Biodiversity, Habitats & Land Use	`BiodivLU`
`E/NatCap`	Upstream Raw-materials Sourcing	`RawSrc`
`E/NatCap`	Resource Use & Material Efficiency	`ResUseMatEff`
`E/NatCap`	Supplier Environmental Due Diligence	`SuppEnvDD`
`E/NatCap`	Water Scarcity & Stress	`WaterSS`
`E/PolWasCiru`	End-of-life Electronics & E-waste	`EWaste`
`E/PolWasCiru`	Packaging Materials & Waste Streams	`PackWaste`
`E/PolWasCiru`	Environmental Compliance & Management Systems	`EnvComp`
`E/PolWasCiru`	Hazardous Emissions & Wastes	`HazWaste`
`E/SustSolTech`	Clean Technology Solutions	`CleanTech`
`E/SustSolTech`	Low-impact & Efficient Buildings	`GreenBldg`
`E/SustSolTech`	Renewable Energy Assets	`ReEneg`
`S/HumCap`	Occupational Health & Safety	`OHS`
`S/HumCap`	Employee Learning & Development	`LearnDev`
`S/HumCap`	Workforce Management & Relations	`WorkManRel`
`S/HumCap`	Job Quality & Working Conditions	`JobQual`
`S/HumCap`	Diversity, Equity & Inclusion	`DEI`
`S/HumCap`	Freedom of Association & Collective Bargaining	`FoACollB`
`S/HumCap`	Local Hiring & Market Presence	`LocHirMarPre`
`S/HumCap`	Labor Standards in Supply Chains	`LabStanSC`
`S/ProdRespCustSafe`	Hazardous Substances & Chemical Safety	`ChemSafe`
`S/ProdRespCustSafe`	Consumer Protection in Finance	`FinProt`
`S/ProdRespCustSafe`	Data Protection & Cybersecurity	`DataSec`
`S/ProdRespCustSafe`	Product Safety, Quality & Reliability	`ProdQual`
`S/ProdRespCustSafe`	Health & Demographic Vulnerability	`HealVuln`
`S/ProdRespCustSafe`	Responsible Marketing & Product Information	`RespMkt`
`S/ProdRespCustSafe`	Responsible & Sustainable Investment	`RSInv`

 doi:10.20944/preprints202602.1970.v2

Table A4. *Cont.*

Pillar/Theme	Focus	Focus abbr.
S/ComRigRisks	Community Impacts & Relations	ComRel
S/ComRigRisks	Human Rights Due Diligence	HRDD
S/ComRigRisks	High-risk & Controversial Sourcing	RiskSrc
S/IncSolSocAcc	Access to Financial Services	FinAcc
S/IncSolSocAcc	Access to Healthcare & Medicines	HealAcc
S/IncSolSocAcc	Digital Connectivity & Inclusion	DigIncl
S/IncSolSocAcc	Nutrition, Diet & Health Opportunities	NutriHeal
G/CorpGov	Board Structure & Oversight	BoardOvr
G/CorpGov	Executive Pay & Incentives	ExecPay
G/CorpGov	Ownership Structures & Control Rights	OwnCtrl
G/CorpGov	Financial Reporting & Accounting Quality	FinRepQ
G/CorpCondInt	Business Ethics & Anti-corruption	EthicsAC
G/CorpCondInt	Tax Strategy, Transparency & Compliance	TaxTrans
G/CorpCondInt	Public Policy & Political Engagement	PolEngage
G/CorpCondInt	Regulatory & Socioeconomic Compliance	RegComp
G/CorpCondInt	Responsible & Local Procurement	RespProc
G/CorpCondInt	Anti-competitive Practices & Fair Competition	FairComp
G/CorpCondInt	Whistleblower Protection & Speak-up Culture	Whistle
G/CorpCondInt	Digital Responsibility, Data & AI Ethics	AIEthics
H/ESGIntAss	Overall ESG Performance & Ratings	ESGRate
H/ESGIntAss	ESG Data Integration & Aggregation	ESGData
H/ESGIntAss	Economic Outcomes & Indirect Impacts	EconImp
H/ESGIntAss	Materiality & Dual-materiality Analysis	DualMat
H/ESGIntAss	ESG Risk & Scenario Analysis	ESGScen
H/ESGIntAss	ESG Reporting & Disclosures	ESGRept
H/ESGIntAss	Greenwashing & Greenhushing	GreenWH
H/ESGIntAss	ESG Controversies & Incident Monitoring	ESGIncid
H/SDFrame	UN SDGs & 2030 Agenda	SDGs
H/SDFrame	Impact-focused Investment	ImpInv
H/SDFrame	Sustainability Reporting Standards (e.g. GRI/SASB)	ReptStd
H/SDFrame	ESG Taxonomies & Standards Mapping	TaxMap
H/SysSustMod	Circular Economy & Resource Loops	CircLoop
H/SysSustMod	End-to-end Supply Chain Sustainability	SCSust
H/SysSustMod	Just Transition & Socio-ecological Change	JustTrans

Appendix D. Construction of the KG4ESG Corpus

This section expands the survey methodology in Section 2 and documents how we curated the KG4ESG corpus. The focus is on *corpus construction*: defining an ESG-aware search space, designing query patterns, harvesting candidate records, and applying PRISMA-style screening and deduplication to obtain the final set of selected works that underpins all analyses in the main text. Figure A3 summarizes this workflow.

 doi:10.20944/preprints202602.1970.v2

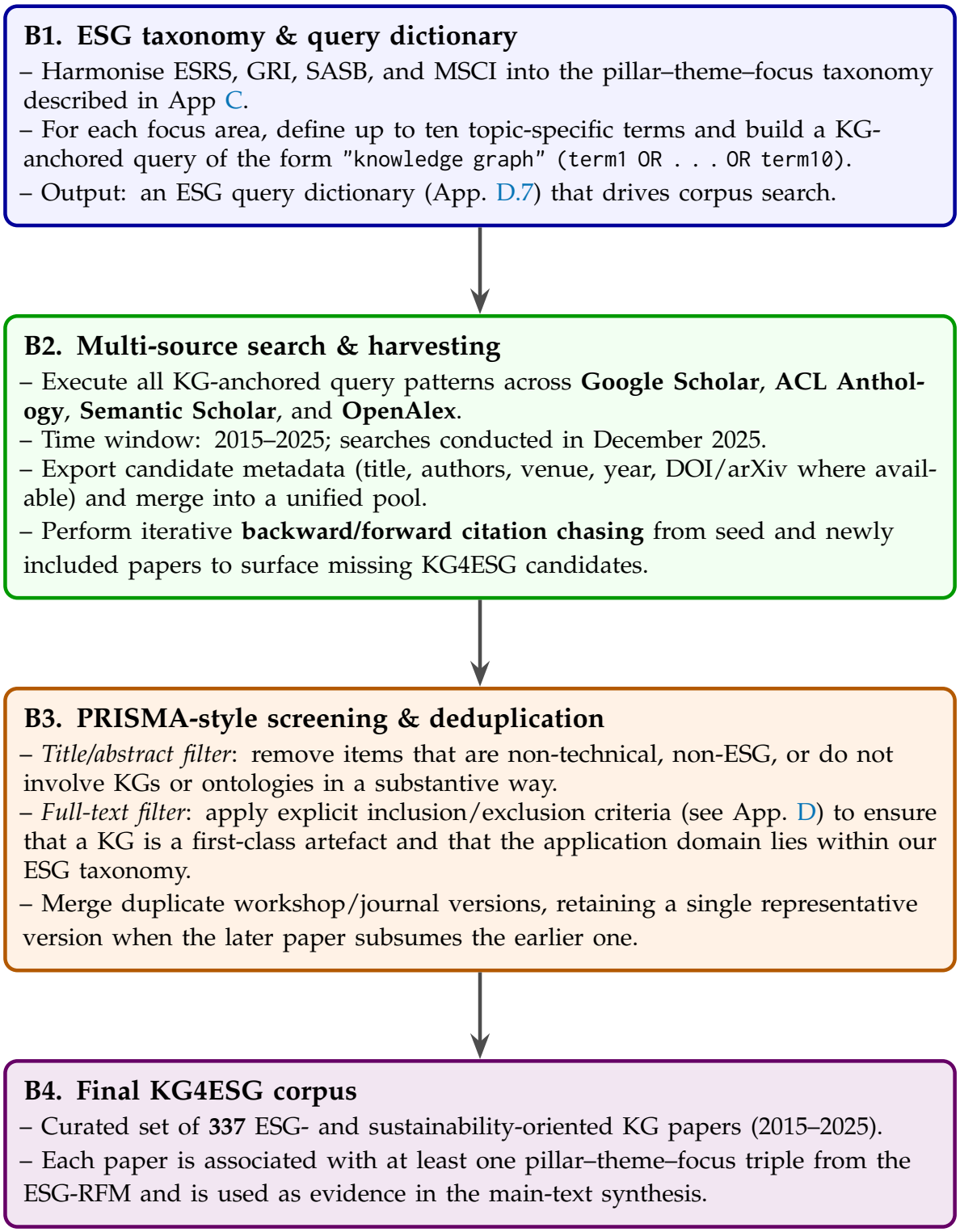

Figure A3. Curation workflow for the KG4ESG corpus. We start from the ESG taxonomy and query dictionary, harvest candidate papers using KG-anchored queries, and apply PRISMA-style screening and deduplication to arrive at a curated corpus of 337 ESG KG papers. Subsequent coding and quantitative meta-analysis are performed on top of this curated corpus and are described elsewhere in the paper.

Appendix D.1. Query Construction & Per–Focus Queries

For each focus in the ESG-RFM pillar–theme–focus hierarchy we construct *two* families of search queries:

- a **KG-anchored query**, which requires the phrase "`knowledge graph`" and combines it with up to ten topic-specific terms; and

 doi:10.20944/preprints202602.1970.v2

- a **background query**, which uses the same topic-specific terms but *omits* the "`knowledge graph`" anchor, so that it reflects the broader ESG/sustainability literature on that focus.

The exact KG-anchored query strings appear in the *Full Google Scholar query* column of Table A5. We run the background queries during corpus harvesting, but we do not report Google Scholar hit counts in the table because they are noisy and time-dependent.

Finally, the **Selected Works** column lists one or more representative KGs from our curated corpus that we judged to be closely aligned with the corresponding focus. These examples are obtained via manual inspection of titles, abstracts, and skimming full texts, not automatically from search results.
Why selected exemplars can surface under multiple foci. Because focus term lists intentionally include synonyms, disclosure hooks, and closely related terminology to increase recall, the same KG paper can be retrieved by multiple focus-level searches. We treat this overlap as a *feature* rather than a defect: it reduces false negatives in ESG, where vocabulary is inconsistent across sectors, jurisdictions, and standards, and where many systems are genuinely multi-topic.

Accordingly, all retrieved candidates are merged and **deduplicated before screening**, and each included work appears only once in the KG4ESG corpus even if it was surfaced by several focus queries. The **Selected Works** column in the query dictionary is therefore *illustrative*: we list each exemplar under the *single focus* that it most directly instantiates based on manual judgement from title/abstract plus full-text inspection. This choice is made for table readability and to provide a clean "most-suitable anchor" for readers, not to claim exclusivity; multi-focus relevance is preserved in our internal coding (and will be retained in the released annotations).

Appendix D.2. Scope & Corpus Construction

We target research whose primary contribution is to *construct, extend, or apply* a knowledge graph (or closely related ontology-style graph) for ESG- or sustainability-related tasks. We accept both academic and industrial papers, including system descriptions, provided that the KG is a first-class artefact rather than a minor internal component.

Starting from the union of all KG-anchored query results and additional manual venue searches (e.g., ACL-like venues, environmental informatics, sustainability and climate conferences, finance and accounting venues), we harvest candidate papers and apply the screening steps shown in Figure A3. After deduplication of workshop and journal versions, this yields the **337** papers that form the KG4ESG corpus used throughout the survey.

Appendix D.3. Search Sources & Workflow

The query dictionary in Appendix D.7 encodes the search patterns used for corpus construction. Each row corresponds to one focus area in the ESG-RFM and provides a KG-anchored query of the form "`knowledge graph`" `(term1 OR . . . OR term10)`. We execute these queries across **four complementary sources**: Google Scholar[5] (broad recall), ACL Anthology[6] (NLP venue completeness), and Semantic Scholar[7] + OpenAlex[8] (structured metadata and citation-graph support). We additionally perform **backward/forward citation chasing** from seed and newly included papers to reduce keyword and indexing bias. All candidates are merged and deduplicated before PRISMA-style screening.
Primary aggregator (Google Scholar). We use Google Scholar as a first-stage aggregator because of its broad, cross-disciplinary coverage and documented high recall relative to single-publisher databases Gusenbauer (2019); Bramer et al. (2018). At the same time, relevance ranking and metadata quality can be noisy Gusenbauer (2019); Bramer et al. (2018). We mitigate this by (i) deriving structured, taxonomy-grounded queries (Appendix C and Appendix D.7), (ii) explicitly anchoring searches on the phrase "knowledge graph", and (iii) applying PRISMA-style screening with pre-defined criteria.

[5] https://scholar.google.com/
[6] https://aclanthology.org/
[7] https://www.semanticscholar.org/
[8] https://openalex.org/

 doi:10.20944/preprints202602.1970.v2

Targeted NLP venue coverage (ACL Anthology). Because KG4ESG is NLP-centered and many relevant papers appear in ACL-family venues, we additionally run targeted searches in the ACL Anthology to reduce the chance that conference/workshop papers are missed or poorly indexed by general-purpose aggregators. Results are merged into the same candidate pool prior to screening.

Structured metadata and citation graph (Semantic Scholar and OpenAlex). We use Semantic Scholar and OpenAlex to (i) cross-check bibliographic fields (venue, year, identifiers), (ii) support high-precision deduplication via DOI/arXiv/normalized titles, and (iii) enable citation-based expansion and verification. In practice, these databases improve reproducibility and reduce errors caused by inconsistent metadata across engines.

Citation chasing and saturation logic. To reduce keyword and indexing bias, we perform backward (reference list) and forward (citing papers) citation chasing from (i) a seed set of high-centrality ESG/sustainability KGs and (ii) borderline-but-included papers discovered during screening. Citation chasing continues iteratively until newly surfaced candidates are predominantly out-of-scope or duplicates under our criteria.

Deduplication and counting conventions. All retrieved candidates from the four sources are merged prior to screening. We deduplicate using a layered strategy: DOI/arXiv identifiers when available, followed by normalized-title matching with manual verification for collisions (e.g., workshop vs. journal extensions). Query-level hit counts are treated only as coarse indicators of topical volume (rather than stable statistics), consistent with bibliometric practice Gusenbauer (2019); Moher et al. (2010). Our overall workflow follows the logic of systematic reviews and meta-analyses Friede et al. (2015); Moher et al. (2010), adapted to the evolving KG/NLP landscape in which many contributions explicitly self-identify with the "knowledge graph" label.

Appendix D.4. Inclusion & Exclusion Criteria

We include a paper in the KG4ESG corpus only if it satisfies all of the following:

1. It defines, instantiates, or operationalises a knowledge graph or ontology that is used as a *central* computational artefact (e.g., for retrieval, reasoning, recommendation, or downstream analytics).
2. Its application domain lies within the ESG and sustainability topics delineated by our taxonomy (i.e., we can assign at least one pillar–theme–focus label).
3. It contains a non-trivial text-centric component (e.g., extraction from reports, scientific articles, regulations, news, or user-generated content), even if the overall system is multimodal.

We exclude works that:

- do not use KGs or ontologies in a substantive way (e.g., purely vector-space or tabular models);
- conduct general ESG text mining (topic modelling, sentiment analysis, etc.) without a KG layer;
- focus exclusively on bibliometric analysis of publications rather than on domain-level ESG KGs;
- operate far outside the corporate-oriented ESG space (e.g., purely scientific climate or ecological models) even if they use KGs, while still acknowledging that insights from those communities are highly relevant.

Appendix D.5. Threats to Validity

The curation process introduces several sources of bias:

- **Search bias.** Anchoring queries on "`knowledge graph`" risks missing ontology- or graph-based work that does not use that exact phrase. We mitigate this by complementing Google Scholar with ACL Anthology (NLP venue coverage), Semantic Scholar and OpenAlex (structured metadata and citation graph), and iterative backward/forward citation chasing from prominent KGs and newly included papers.
- **Hit-count noise.** Google Scholar hit counts are approximate and time-dependent. The volume labels in Table A5 should be read as coarse activity indicators, not as precise bibliometric measures.

 doi:10.20944/preprints202602.1970.v2

Appendix D.6. Limitations

The KG4ESG corpus should be interpreted as a carefully curated but *non-exhaustive* sample of ESG and sustainability KGs. It is shaped by our taxonomy, our query patterns, and our inclusion criteria. In particular:

- some adjacent graph or ontology work may be missing because it does not self-identify as a "knowledge graph";
- fast-moving areas (e.g., LLM- and agentic KG workflows) may be under-represented in the most recent years;
- borderline cases between pillars or themes may reasonably admit alternative classifications.

We therefore release the taxonomy and query dictionary as reusable artefacts so that future work can extend, refine, or contest the KG4ESG corpus as the literature evolves.

Appendix D.7. KG4ESG Corpus Query Dictionary

For every focus in ESG-RFM, we define a small query dictionary used to retrieve candidate literature.

KG-anchored query construction. Because this survey targets knowledge graph methods, every retrieval query is anchored with the fixed phrase "`knowledge graph`". The table below reports only the focus-specific term list (an OR-list). On engines that accept multi-term keyword queries (Google Scholar, ACL Anthology, OpenAlex), the executable KG-anchored query is instantiated by prepending the anchor phrase and grouping the terms with OR. Concretely:

"`knowledge graph`" ($term_1$ `OR` $term_2$ `OR` . . . `OR` $term_k$), $k \leq 10$.

For example, the *Greenhouse Gas Emissions* focus instantiates:

"`knowledge graph`" ("`GHG emissions`" `OR` "`CO2 emissions`" `OR` . . . `OR` "`Scope 3 emissions`").

All searches were executed in December 2025.

Two query patterns and year-wise execution. For each focus area, we execute two complementary query patterns: (i) a KG-anchored query that explicitly includes the "`knowledge graph`" constraint, and (ii) a **background query** that omits this anchor while reusing the same focus-specific term list. For Google Scholar and OpenAlex, both query patterns are executed *year-wise* for each year from 2015 to 2025 using the corresponding year filters provided by the platforms. For ACL Anthology, due to search-engine limitations that do not support reliable year-level filtering, each query pattern is executed *without explicit year constraints*, and publication years are subsequently resolved from the retrieved metadata during corpus curation.

Semantic Scholar execution (OR-by-union). Semantic Scholar's public web search does not support Boolean operators (e.g., OR-lists), although it supports quoted phrases.[9] Therefore, the OR-list cannot be issued as a single query. Instead, we **decompose** each focus-year query into k single-term sub-queries (with phrases quoted) and then **combine** the results by set union:

"`knowledge graph`" "`term`$_1$", "`knowledge graph`" "`term`$_2$", . . . , "`knowledge graph`" "`term`$_k$".

We apply the same decomposition for background retrieval (dropping the anchor and querying "`term`$_i$").

Why up to ten terms per focus? We cap each focus term list at **ten** terms ($k \leq 10$) as a practical compromise between (i) **recall** (covering common synonyms, abbreviations, metric/reporting language, and standard-specific hooks) and (ii) **portability and auditability** (keeping queries short enough to run consistently across engines and to inspect/maintain as a curated dictionary). Term lists are derived from the focus definition and the MSCI/GRI/ESRS/SASB crosswalk vocabulary, then augmented with frequently used synonyms and abbreviations observed in pilot searches and seed papers. The cap is

[9] https://www.semanticscholar.org/faq/boolean-queries

 doi:10.20944/preprints202602.1970.v2

also important for Semantic Scholar: because OR must be implemented by decomposition, k directly controls the number of required sub-queries.

Illustration (coverage-oriented term selection). To increase topical coverage beyond a single keyword, term sets intentionally mix (i) *phenomenon-level* language and (ii) *disclosure / reporting* language. For instance, the *Physical Climate Risk & Vulnerability* focus includes both risk/hazard terms (e.g., "`physical climate risk`", "`climate vulnerability`") and disclosure hooks (e.g., "`TCFD reporting`", "`scenario analysis`"), while *Biodiversity, Habitats & Land Use* combines ecological processes (e.g., "`deforestation`", "`land use change`") with emerging standards vocabulary (e.g., "`TNFD`"). Despite these safeguards, no fixed dictionary can be exhaustive in ESG: terminology is sector-, jurisdiction-, and time-dependent. We therefore treat the query dictionary as a recall-oriented scaffold and further mitigate residual keyword/indexing bias through backward/forward citation chasing and iterative screening refinement (Appendix D.3).

Each row in the table below corresponds to one focus area and contains the following fields:

- **Pillar / Theme / Focus** — the position of the focus in the ESG-RFM pillar–theme–focus hierarchy.
- **Searched Terms** — the focus-specific term list of up to ten phrases used to instantiate the KG-anchored query (and the corresponding background query).
- **Selected Works** — one or more representative KG4ESG papers from our curated corpus that directly instantiate or strongly target the focus. These examples are selected manually via full-text inspection.

Table A5. KG4ESG search-term dictionary across pillars, themes, and focus areas (all searches executed in December 2025).

Pillar	Theme	Focus	Search Query Terms	Selected Works
E	Clim Chg	GHG	"GHG emissions" | "CO2 emissions" | "carbon emissions" | "carbon intensity" | "emissions inventory" | "emissions factor" | "carbon disclosure" | "Scope 1 emissions" | "Scope 2 emissions" | "Scope 3 emissions"	Fu et al. (2025); Ji et al. (2025); Sheng et al. (2025); CarbonKG; NW1; Machine Knowledge Graph (MKG); Zhao (2023); Wu et al. (2023); Oladeji and Mousavi (2023); Ma et al. (2023); E-Liability Knowledge Graph; Emission Conversion Factors Knowledge Graph (ECF KG / CFKG); Sharma et al. (2022)
E	Clim Chg	Phys Clim Risk	"physical climate risk" | "climate resilience" | "climate vulnerability" | "climate risk assessment" | "climate adaptation" | "climate transition risk" | "acute climate risk" | "TCFD reporting" | "climate scenario analysis" | "climate risk disclosure"	Wang et al. (2025); ClimaFactsKG; Stade et al. (2025); Aivalis et al. (2025); Remote Sensing Early Warning Knowledge Graph (RSEW-KG); Wu et al. (2024); Hofmeister et al. (2024); Androna et al. (2024); Extreme Climate Architecture Knowledge Graph; SILVANUS Knowledge Graph; I-KNOW-FOO knowledge graph; KnowUREnvironment; Mishra and Mittal (2021)
E	Clim Chg	Clim Fin	"financing environmental impact" | "sustainable finance" | "green finance" | "transition finance" | "coal financing" | "green bonds" | "ESG lending" | "sustainability-linked loans" | "climate finance" | "green investment banking"	Ushio et al. (2025); NatureKG; Vasiliu et al. (2024); Peng et al. (2023); ESG Knowledge Graph
E	Clim Chg	Energy Mix	"energy use" | "energy consumption" | "energy efficiency" | "energy intensity" | "fuel mix" | "electricity mix" | "renewable energy consumption" | "electricity consumption" | "energy management system" | "energy performance"	Energy Knowledge Graph (EKG); Karjou et al. (2025); Hanžel et al. (2025); OfficeGraph; Wang et al. (2023); Janev et al. (2022); Li et al. (2022); Aryan et al. (2021); Xiaoping et al. (2021); Janev et al. (2021); Wang et al. (2021); Power Knowledge Graph; Zhao et al. (2020); Tang et al. (2019)
E	Clim Chg	Prod Foot	"product carbon footprint" | "product GHG accounting" | "product life cycle emissions" | "life cycle assessment" | "LCA" | "product environmental footprint" | "product carbon intensity" | "ISO 14067" | "product carbon label" | "product-level emissions disclosure"	LCA Knowledge Graph; LCI knowledge graph; ForestFoodKG-KG; Asset Life Cycle Knowledge Graph (ALC KG); Greif et al. (2024); Peng et al. (2024); Li et al. (2024); LCIKG; Wang et al. (2022)
E	Nat Cap	Biodiv LU	"biodiversity loss" | "biodiversity conservation" | "land use change" | "deforestation" | "habitat conversion" | "ecosystem services" | "nature-related risk" | "nature loss" | "TNFD" | "biodiversity impact assessment"	CropDP-KG; Tabanao et al. (2025); Firmographica; FooDS; GeoKG (Habitat GeoKG); SOCKG; ECOLOPES KG; METRIN-KG; KG-PLUB; Ashworth et al. (2025); Semantic Knowledge Graph of European Mountain Value Chains; Wen et al. (2024); Nature FIRST KG; Le Guillarme and Thuiller (2023); OpenBiodiv biodiversity knowledge graph; Biospytial Knowledge Graph
E	Nat Cap	Raw Src	"raw material sourcing" | "raw material supply" | "responsible raw materials" | "origin of raw materials" | "upstream commodity sourcing" | "raw material traceability" | "mineral sourcing" | "agricultural commodity sourcing" | "certified raw materials" | "conflict-free sourcing"	Jin et al. (2025); Liu et al. (2023); Ameri et al. (2023)

 doi:10.20944/preprints202602.1970.v2

Table A5. *Cont.*

Pillar	Theme	Focus	Search Query Terms	Selected Works
E	Nat Cap	Res Use Mat Eff	"materials use" | "material consumption" | "resource efficiency" | "material efficiency" | "resource productivity" | "material intensity" | "resource use reduction" | "use of recycled materials" | "recycled content" | "resource conservation"	Chen et al. (2025)
E	Nat Cap	Supp Env DD	"supplier environmental assessment" | "supplier environmental performance" | "environmental supplier audit" | "environmental supplier evaluation" | "environmental supplier questionnaire" | "green supplier selection" | "supplier environmental scorecard" | "supplier environmental rating" | "environmental supply chain assessment" | "environmental supplier monitoring"	N.A.
E	Nat Cap	Water SS	"water scarcity" | "water stress" | "water risk" | "water footprint" | "basin water risk" | "water withdrawal" | "water consumption" | "water availability" | "water use efficiency" | "water stewardship"	WHOW-KG; Wang et al. (2025); Papageorgiou et al. (2025); Yang et al. (2025); Tang et al. (2024); Wang et al. (2024); Sun et al. (2024); Papageorgiou et al. (2023); Ospan et al. (2023); He et al. (2023); Yan et al. (2023); Wang et al. (2023a); Wang et al. (2023); Wang et al. (2023b); Water Knowledge Graph (WKG) and Water Information Network (WIN); Liang et al. (2022); Water Affair Knowledge Graph
E	Pol Was Ciru	EWaste	"electronic waste" | "e-waste" | "WEEE" | "end-of-life electronics" | "electronic waste recycling" | "e-waste management" | "e-waste collection" | "e-waste disposal" | "informal e-waste sector" | "e-waste policy"	Hou et al. (2025); Wang et al. (2025); Yin et al. (2025); Wu et al. (2024); Global EEE Green Design Knowledge Graph; Wang et al. (2023)
E	Pol Was Ciru	Pack Waste	"plastic packaging" | "single-use packaging" | "single-use plastics" | "packaging waste" | "packaging circularity" | "recyclable packaging" | "packaging design for recycling" | "packaging recyclability" | "packaging extended producer responsibility" | "sustainable packaging design"	Thapa et al. (2025); Aprilia et al. (2023)
E	Pol Was Ciru	Env Comp	"environmental compliance" | "environmental regulation compliance" | "environmental fines" | "environmental violations" | "environmental management system" | "EMS certification" | "environmental policy" | "environmental audit" | "environmental legal compliance" | "environmental enforcement"	KnowWhereGraph (KWG); Yang et al. (2025); Knowledge Graph of Dangerous Goods (KGDG); Zhang et al. (2021); Wang et al. (2021)
E	Pol Was Ciru	Haz Waste	"hazardous waste" | "toxic releases" | "toxic emissions" | "industrial air pollution" | "industrial water pollution" | "soil contamination" | "hazardous substances" | "toxic waste management" | "industrial waste management" | "pollutant discharge"	Zheng et al. (2025); Cheng et al. (2025); Katzenstein and Etcheverry (2025); Du et al. (2024); Han et al. (2022)
E	Sust Sol Tech	Clean Tech	"clean technology" | "cleantech innovation" | "low-carbon technology" | "industrial decarbonization technology" | "energy-efficient technologies" | "emissions-reduction technologies" | "green technology innovation" | "climate technology solutions" | "clean technology adoption" | "industrial clean tech"	Stagnol et al. (2023)
E	Sust Sol Tech	Green Bldg	"green building" | "sustainable building" | "energy-efficient buildings" | "building energy performance" | "LEED certification" | "BREEAM certification" | "zero-energy buildings" | "net zero buildings" | "green building certification" | "green building design"	Li et al. (2024); Wang et al. (2024); Quek et al. (2024); PGD-KG; Delgoshaei et al. (2022)
E	Sust Sol Tech	Re Eneg	"renewable energy" | "solar power" | "solar photovoltaics" | "wind power" | "onshore wind power" | "offshore wind power" | "hydropower projects" | "renewable energy projects" | "renewable energy investment" | "corporate renewable sourcing"	knowledge graph structure; Multi-modal Process Knowledge Graph for Wind Turbines (MPKG-WT); Pang et al. (2024); Energy Knowledge Graph; XAI4Wind Knowledge Graph; Energy Knowledge Graph (EKG); Energy Knowledge Graph (EKG) AEC-KG; FFHKG; Zhang et al. (2025); Shi and Wu (2024); Li et al.; Jiang et al. (2024); Zhou et al. (2024); Wen (2024); MAKG; Peng et al. (2023); Construction Accident Knowledge Graph (CAKG); Wang and El-Gohary (2023); Simone et al. (2023); Job Hazard Analysis Knowledge Graph (JHAKG); Risk Knowledge Graph in Railway Safety (RKGRS); Pedro et al. (2022); Mao et al. (2020); Fang et al. (2020)
S	Hum Cap	OHS	"occupational health and safety" | "workplace health and safety" | "workplace injuries" | "safety incidents" | "lost time injury" | "lost time injury frequency rate" | "safety management systems" | "process safety management" | "safety culture" | "occupational safety"	
S	Hum Cap	Learn Dev	"human capital development" | "employee development" | "learning and development" | "upskilling" | "reskilling" | "talent development" | "workforce development" | "employee training" | "skills development" | "continuous learning culture"	Person–Job Temporal Knowledge Graph; Fathi (2024); MetaKG; Loh et al. (2022); Weichselbraun et al. (2022); Skills & Occupation Knowledge Graph
S	Hum Cap	Work Man Rel	"labor management" | "employee relations" | "workforce relations" | "employee relations strategy" | "employee engagement" | "labor dispute resolution" | "employee voice mechanisms" | "workplace dialogue" | "labor-management cooperation" | "employee relations climate"	JobEdKG and T-JobEdKG; KG-IRDM; Al Akasheh et al. (2024)
S	Hum Cap	Job Qual	"employment quality" | "decent work" | "fair wages" | "living wage" | "working conditions" | "job security" | "precarious employment" | "working time" | "employee benefits" | "employment contracts"	Human-Centered Knowledge Graph (HCKG); Salary Knowledge Graph; Wang et al. (2022)
S	Hum Cap	DEI	"workplace diversity" | "equal opportunity" | "non-discrimination" | "workforce diversity" | "gender diversity" | "racial diversity" | "pay equity" | "inclusive workplace" | "diversity and inclusion" | "DEI program"	Aravind Krishnan and Deepak (2023)

 doi:10.20944/preprints202602.1970.v2

Table A5. *Cont.*

Pillar	Theme	Focus	Search Query Terms	Selected Works
S	Hum Cap	FoA Coll B	"freedom of association" | "collective bargaining" | "trade union rights" | "unionization" | "collective labor agreements" | "collective agreements coverage" | "social dialogue" | "works council" | "employee representation" | "union density"	N.A.
S	Hum Cap	Loc Hir Mar Pre	"market presence" | "local hiring" | "local employment" | "local recruitment" | "local wage levels" | "local management hiring" | "local content in employment" | "local workforce development" | "local economic integration" | "local talent pipeline"	N.A.
S	Hum Cap	Lab Stan SC	"supplier labor standards" | "forced labor" | "child labor" | "modern slavery" | "decent work in supply chains" | "supply chain human rights" | "labor exploitation in supply chains" | "supply chain due diligence" | "ethical supply chains" | "labor rights in supply chains"	Zheng and Brintrup (2025); Verona et al. (2018)
S	Prod Resp Cust Safe	Chem Safe	"chemical safety" | "hazardous substances" | "toxic chemicals" | "substance restrictions" | "REACH regulation" | "chemical risk management" | "chemical hazard assessment" | "chemical exposure limits" | "hazard communication" | "chemical safety regulation"	CPSKG; Hazardous Chemical Accident Knowledge Graph (HCAKG); Multimodal Knowledge Graph (MMKG); Da Silveira et al. (2024); TERA; SEARCH-KG; Zheng et al. (2021)
S	Prod Resp Cust Safe	Fin Prot	"consumer financial protection" | "financial consumer protection" | "fair lending" | "predatory lending" | "consumer credit risk" | "responsible lending" | "over-indebtedness" | "consumer protection regulation" | "financial conduct regulation" | "consumer financial rights"	N.A.
S	Prod Resp Cust Safe	Data Sec	"data privacy" | "information privacy" | "cybersecurity" | "information security" | "personal data protection" | "data protection regulation" | "GDPR compliance" | "data breach" | "privacy by design" | "data security management"	Liu et al. (2025); Sharma et al. (2025); Cheng et al. (2025); IoT-Reg Knowledge Graph; Cotti et al. (2025); Qiao et al. (2025); AttacKG+; Hu et al. (2024); PrivComp-KG; VulKG; Gilliard et al. (2024); Kurniawan et al. (2024); Li et al. (2023); Liu and Zhan (2023); IoT-Reg Knowledge Graph; PRIVAFRAME; AttacKG; CSKG4APT; Open-CyKG; Cybersecurity Knowledge Graph (CSKG); SEPSES Cybersecurity Knowledge Graph; Elluri et al. (2018); Jia et al. (2018)
S	Prod Resp Cust Safe	Prod Qual	"product safety" | "product safety incidents" | "product recalls" | "safety defects" | "product quality" | "product quality assurance" | "product reliability" | "consumer product safety" | "product safety regulation" | "product conformity assessment"	Huang and Song (2025); ESKG; MPKG; QChsG; Jing and Li (2024); Causal Quality-related Knowledge Graph (CQKG); Ren et al. (2024); Human–Cyber–Physical Knowledge Graph; Yang et al. (2022); Wang et al. (2021); Liu et al. (2021); Zhang et al. (2021); Sun and Wang (2019); Manufacturing Knowledge Graph (MKG)
S	Prod Resp Cust Safe	Heal Vuln	"health and demographic risk" | "demographic risk" | "aging population risk" | "population aging" | "epidemiological transition" | "chronic disease burden" | "pandemic risk" | "population health trends" | "demographic change" | "health risk exposure"	Elderly Advantages Knowledge Graph; Li et al. (2024); Duan et al. (2018)
S	Prod Resp Cust Safe	Resp Mkt	"responsible marketing" | "responsible advertising" | "product labeling" | "product labeling" | "marketing ethics" | "misleading advertising" | "green marketing claims" | "sustainability claims" | "fair marketing practices" | "truthful product information"	N.A.
S	Prod Resp Cust Safe	RSInv	"responsible investment" | "responsible investing" | "sustainable investing" | "ESG investing" | "ESG integration" | "ESG screening" | "stewardship investing" | "shareholder engagement" | "active ownership" | "fiduciary duty and ESG"	Social-Impact Funding Knowledge Graph
S	Com Rig Risks	Com Rel	"community engagement" | "stakeholder engagement" | "social license to operate" | "indigenous rights" | "community impact" | "local community opposition" | "community consultation" | "community grievance mechanisms" | "community development programs" | "community benefit agreements"	Gordon et al. (2021)
S	Com Rig Risks	HRDD	"human rights" | "human rights impact assessment" | "HRIA" | "salient human rights issues" | "UN Guiding Principles" | "UNGP implementation" | "OECD due diligence" | "corporate human rights assessment" | "human rights risk assessment" | "human rights grievance mechanism"	PREJUST4WOMAN; Observatory Knowledge Graph (OKG)
S	Com Rig Risks	Risk Src	"controversial sourcing" | "conflict minerals" | "responsible mineral sourcing" | "cobalt supply chain" | "palm oil sourcing" | "high-risk sourcing countries" | "responsible mineral supply chains" | "ethical mineral sourcing" | "responsible cobalt sourcing" | "responsible palm oil"	Supply Chain Knowledge Graph (SC-KG)
S	Inc Sol Soc Acc	Fin Acc	"access to finance" | "access to financial services" | "financial inclusion" | "microfinance" | "digital financial services" | "inclusive finance" | "unbanked populations" | "underserved customers" | "mobile money" | "SME access to finance"	Mitra et al. (2024); Alam and Ali (2022); Beydon et al. (2020)
S	Inc Sol Soc Acc	Heal Acc	"access to health care" | "access to medicines" | "affordable medicines" | "healthcare affordability" | "telemedicine services" | "healthcare availability" | "universal health coverage" | "healthcare equity" | "rural health care access" | "primary health care access"	HealthEQKG; Uddin et al. (2025); NHANES Knowledge Graph; Diseasomics knowledge graph; Huang et al. (2019)

 doi:10.20944/preprints202602.1970.v2

Table A5. *Cont.*

Pillar	Theme	Focus	Search Query Terms	Selected Works
S	Inc Sol Soc Acc	Dig Incl	"access to communications" | "digital inclusion" | "digital divide" | "access to internet" | "broadband access" | "mobile connectivity" | "ICT access" | "affordable connectivity" | "digital infrastructure" | "telecommunications access"	Jain et al. (2025); Xiao and Sun (2023)
S	Inc Sol Soc Acc	Nutri Heal	"nutrition and health" | "healthy foods" | "fortified foods" | "nutrition interventions" | "obesity prevention" | "diet-related health" | "healthy diet" | "nutrition security" | "functional foods" | "nutrition labeling"	Wu et al. (2025); Gao et al. (2025); NRKG; Food4healthKG; GENA; KG4NH; KG4NH; FoodKG; Chi et al. (2018)
G	Corp Gov	Board Ovr	"corporate board" | "board independence" | "independent directors" | "board diversity" | "board gender diversity" | "board oversight" | "board composition" | "director overboarding" | "board refreshment" | "board tenure"	N.A.
G	Corp Gov	Exec Pay	"executive compensation" | "CEO compensation" | "CEO salary" | "say on pay" | "CEO pay ratio" | "incentive alignment" | "performance-based pay" | "remuneration policy" | "compensation committee" | "executive pay governance"	N.A.
G	Corp Gov	Own Ctrl	"ownership and control" | "dual-class shares" | "pyramidal ownership" | "control rights" | "controlling shareholders" | "ownership concentration" | "ownership structure" | "family ownership" | "blockholder ownership" | "minority shareholder protection"	Equity Knowledge Graph (EKG); LCAIM Knowledge Graph; Company Knowledge Graph; Atzeni et al. (2020)
G	Corp Gov	Fin RepQ	"accounting irregularities" | "financial restatement" | "earnings management" | "aggressive accounting" | "audit quality" | "financial reporting quality" | "accounting fraud" | "financial statement fraud" | "creative accounting" | "accounting enforcement"	Financial Knowledge Graph (FKG); Badmus et al. (2025); Zhu et al. (2025); Osman et al. (2024); FSFD-TLKG; Supplier–Customer Knowledge Graph; Wu et al. (2022); Manager Knowledge Graph; Financial Knowledge Graph (FKG); FSFD
G	Corp Cond Int	Ethics AC	"business ethics" | "corporate ethics" | "anti-corruption" | "anti-bribery" | "corruption risk" | "integrity management" | "ethics training" | "ethics and compliance program" | "corporate misconduct" | "ethical culture"	Liu et al. (2021); Gao and Liu (2021)
G	Corp Cond Int	Tax Trans	"tax transparency" | "tax disclosure" | "tax avoidance" | "tax planning" | "effective tax rate" | "country-by-country reporting" | "tax policy" | "corporate tax strategy" | "tax justice" | "aggressive tax planning"	Zhang (2025); Tax Law Knowledge Graph (TaxKG); Enterprise Knowledge Graph; TaxGraph – Global Multinational Tax Planning Knowledge Graph
G	Corp Cond Int	Pol Engage	"corporate political activity" | "corporate political engagement" | "lobbying activities" | "public policy advocacy" | "political contributions" | "political donations" | "policy influence" | "trade association lobbying" | "political spending disclosure" | "regulatory lobbying"	N.A.
G	Corp Cond Int	Reg Comp	"socioeconomic compliance" | "regulatory compliance" | "non-financial compliance" | "competition law compliance" | "product regulatory compliance" | "social regulation compliance" | "legal non-compliance" | "fines and sanctions" | "compliance violations" | "regulatory enforcement actions"	GovGraph; Vanapalli et al. (2025); Chung et al. (2025); Jomraj et al. (2025); Boukhelifa and Merabet (2024); Regulatory Knowledge Graph; Tauqeer et al. (2022)
G	Corp Cond Int	Resp Proc	"responsible procurement" | "local procurement" | "local suppliers" | "inclusive procurement" | "responsible purchasing" | "sustainable purchasing policy" | "ethical procurement" | "supplier diversity" | "community-focused procurement" | "procurement sustainability"	N.A.
G	Corp Cond Int	Fair Comp	"anti-competitive behavior" | "anti-competitive behavior" | "competition law" | "antitrust violations" | "cartel" | "price fixing" | "abuse of dominance" | "market collusion" | "bid rigging" | "fair competition"	Tang and Chi (2025)
G	Corp Cond Int	Whistle	"whistleblower protection" | "whistleblowing system" | "speak-up culture" | "whistleblowing hotline" | "retaliation against whistleblowers" | "internal reporting channels" | "ethics reporting" | "whistleblower policy" | "whistleblowing in organizations" | "speak-up mechanism"	N.A.
G	Corp Cond Int	AI Ethics	"AI ethics" | "algorithmic accountability" | "algorithmic fairness" | "responsible AI" | "digital responsibility" | "AI governance" | "data ethics" | "ethical AI principles" | "AI risk management" | "trustworthy AI"	TAIR (Trustworthy AI Requirements) knowledge graph
H	ESG Int Ass	ESG Rate	"ESG performance" | "ESG rating" | "overall ESG score" | "ESG benchmarking" | "sustainability score" | "ESG assessment" | "ESG index rating" | "issuer ESG rating" | "portfolio ESG rating" | "ESG ranking"	Cai et al. (2025); Cai et al. (2025); Wu et al. (2025); ESG-KG; Oksanen et al. (2022)
H	ESG Int Ass	ESG Data	"ESG data integration" | "ESG data aggregation" | "ESG data harmonization" | "sustainability data platforms" | "ESG data pipelines" | "ESG data quality" | "ESG datasets" | "ESG data sources" | "ESG database integration" | "ESG data interoperability"	Hassan Nassar et al. (2025)
H	ESG Int Ass	Econ Imp	"economic performance" | "economic value generated" | "economic value distributed" | "indirect economic impacts" | "local economic impact" | "inclusive economic growth" | "economic development impact" | "social return on investment" | "socioeconomic impact assessment" | "economic contribution analysis"	Spatiotemporal Knowledge Graph (STKG)

 doi:10.20944/preprints202602.1970.v2

Table A5. *Cont.*

Pillar	Theme	Focus	Search Query Terms	Selected Works
H	ESG Int Ass	Dual Mat	"materiality assessment" | "ESG materiality" | "double materiality" | "impact materiality" | "financial materiality" | "sustainability materiality" | "materiality matrix" | "stakeholder materiality assessment" | "ESG topic prioritization" | "material issues identification"	Usmanova and Usbeck (2024)
H	ESG Int Ass	ESG Scen	"ESG risk management" | "sustainability risk modeling" | "climate-related financial risk analysis" | "transition and physical climate risk evaluation" | "ESG scenario-based analytics" | "climate-aligned scenario planning" | "ESG exposure and vulnerability assessment" | "sustainability stress-test analytics" | "climate-financial impact projection" | "climate-ESG risk propagation"	LinkClimate knowledge graph
H	ESG Int Ass	ESG Rept	"ESG reporting" | "sustainability disclosure" | "ESG disclosure" | "sustainability statement" | "ESG transparency" | "corporate sustainability reporting" | "non-financial disclosure" | "integrated ESG reporting" | "ESG narrative reporting" | "ESG metrics disclosure"	Yu et al. (2025); ESGSenticNet; Dolha et al. (2025); Power Equipment Management Using Knowledge Graph; Raikar and Deepak (2024)
H	ESG Int Ass	Green WH	"greenwashing" | "green hushing" | "greenhushing" | "misleading environmental claims" | "misleading sustainability claims" | "green claims" | "environmental claim substantiation" | "sustainability claim verification" | "green claims directive" | "ESG communication integrity"	DeBellis et al. (2025); EmeraldGraph
H	ESG Int Ass	ESG Incid	"ESG controversies" | "ESG controversy" | "sustainability controversies" | "ESG incident database" | "ESG incident detection" | "ESG news monitoring" | "adverse ESG events" | "ESG negative screening" | "controversial business practices" | "controversy scoring"	Iwata et al. (2025)
H	SD Frame	SDGs	"SDGs" | "Sustainable Development Goals" | "Agenda 2030" | "sustainability targets" | "SDG indicators" | "SDG alignment" | "SDG monitoring" | "SDG integration" | "SDG evaluation" | "sustainable development framework"	SDG-KG; Benjira et al. (2025); Regional Graph (RG) + Intent Graph (IG); Lin and Liao (2025); Mishra et al. (2025); Kilanioti and Papadopoulos (2024); Benjira et al. (2024); UrbanKG; SustainGraph; DaanMatch Knowledge Graph (DaanKG); Mandilara et al. (2023); Orellana and Piedra (2021); Eguiguren and Piedra (2020); Eguiguren and Piedra (2019)
H	SD Frame	Imp Inv	"impact investing" | "impact investment funds" | "impact measurement" | "impact financing" | "social impact investing" | "environmental impact investing" | "impact evaluation" | "impact reporting" | "impact fund strategies" | "sustainable impact investing"	N.A.
H	SD Frame	Rept Std	"sustainability reporting" | "GRI standards" | "SASB standards" | "GRI/SASB alignment" | "sustainability reporting framework" | "non-financial reporting standards" | "GRI-compliant reporting" | "SASB-based disclosure" | "sustainability reporting taxonomy" | "ESG reporting standards"	ESG Metric Knowledge Graph (ESGMKG); Bronzini et al. (2024); RSOKG; Zhou et al. (2024); Gupta et al. (2024); Diamantini et al. (2022)
H	SD Frame	Tax Map	"ESG taxonomy" | "sustainable finance taxonomy" | "EU taxonomy" | "green taxonomy" | "ESG classification system" | "ESG standard mapping" | "ESG standards alignment" | "sustainability reporting standards" | "ESG ontology" | "sustainability ontology"	Koloski et al. (2022)
H	Sys Sust Mod	Circ Loop	"circular economy" | "closed-loop supply chains" | "resource circularity" | "material circularity" | "product life extension" | "reuse and remanufacturing" | "circular business models" | "industrial symbiosis" | "circular value chains" | "circular economy metrics"	Zhao et al. (2025); Belova et al. (2025); Housing Passport Knowledge Graph (HPKG); W2RKG; Marconnet et al. (2024); Lu et al. (2024); ScrapKG; Kebede et al. (2023)
H	Sys Sust Mod	SC Sust	"sustainable supply chains" | "end-to-end supply chain sustainability" | "ESG in supply chains" | "supply chain sustainability assessment" | "multi-tier supply chain ESG" | "responsible global supply chains" | "sustainable sourcing and logistics" | "supply chain sustainability metrics" | "supply chain ESG performance" | "holistic supply chain risk"	Felder et al. (2025); AlMahri et al. (2025); Heus et al. (2025); Zheng et al. (2025)
H	Sys Sust Mod	Just Trans	"just transition" | "climate justice" | "energy justice" | "low-carbon transition pathways" | "socially just decarbonization" | "green jobs transition" | "coal phase-out justice" | "inclusive climate transition" | "socio-ecological transformation" | "just transition planning"	GeoOutageKG

Appendix E. Supplementary Analysis for Construction Paradigms

Paradigms are defined by the dominant evidence-to-schema mapping operator rather than by the mere presence of an ontology. P1 denotes *ontology-first lifting and deterministic integration* in which the semantic contract is the primary artifact and population is largely manual or ETL-style; typical exemplars emphasize standards interoperability and metric semantics Diamantini et al. (2022); Yu et al. (2024); Driller and Trang (2024); Koloski et al. (2022). P2 denotes *rule/supervised NLP/ML extraction*

 doi:10.20944/preprints202602.1970.v2

that treats the schema as a label space for extraction and classification Oksanen et al. (2022); Burel and Alani (2025); Blin et al. (2023); Bauer et al. (2022). P3 denotes *LLM-assisted structuring and alignment* in which structured outputs are generated under schema constraints, commonly with retrieval grounding Yu et al. (2025); Bronzini et al. (2024); Zhou et al. (2024); Usmanova and Usbeck (2024); Iwata et al. (2025). P4 denotes *agentic/tool-using pipelines with iterative validation and repair*, where explicit controllers decompose construction or querying into multi-step actions, enabling traceability, constraint checks, and correction loops; this includes both LLM+tool systems for text-to-graph interaction Zhao et al. (2025); Karimanzira et al. (2025) and tool-centric dynamic KG systems without LLMs Hofmeister et al. (2024); Quek et al. (2024); Xie et al. (2024); Ong et al. (2022).

Table A6. Conceptual distinctions among P1–P4 framed around the evidence-to-schema mapping operator and ESG-grade validation requirements.

P	Dominant mapping operator	Primary bottleneck addressed	What is fixed vs. adaptive	Typical failure modes	Validation hooks aligned with ESG use
P1	Schema/ontology engineering + deterministic lifting	Semantic precision, interoperability, provenance design	Schema fixed; mappings largely fixed; updates costly	Coverage limits; slow updates; brittle mappings under drift	Schema scope/versioning; provenance model; constraint completeness; update policy
P2	Supervised/rule-based NLP over a label space	Scalable, repeatable extraction from text	Schema fixed; models learn mapping to types/relations	Label scarcity; domain shift; pipeline compounding; weak alignment	Extraction metrics; error propagation analysis; conformance rate; linking quality
P3	Schema-conditioned generation (often with retrieval)	Semantic elasticity, long-tail coverage, cross-standard alignment	Schema as constrained IR; mapping flexible and context-dependent	Hallucinated triples; inconsistent typing; silent schema drift	Grounding strategy; structured output validation; provenance enforcement; constraint checks
P4	Tool-using controllers with iterative validation/repair	End-to-end reliability under high-stakes workflows	Schema as callable contract; adaptive action sequences	Tool misuse; feedback loops; hidden cascades without governance	Action traces; explicit constraint checks; conflict resolution; escalation protocols

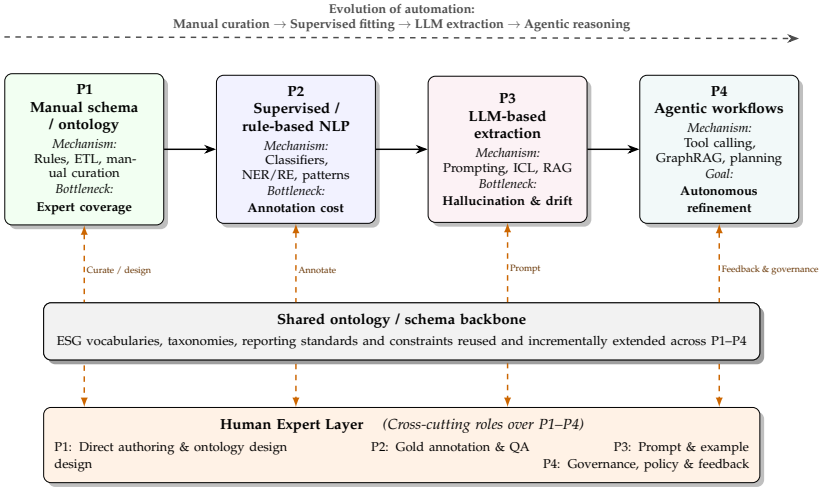

Figure A4. Horizontal evolution of ESG KG construction paradigms (P1–P4). All paradigms assume and reuse a shared ontology/schema backbone (middle), which encodes ESG vocabularies, reporting standards and constraints. The Human Expert Layer (bottom) supports all paradigms, shifting from direct authoring and ontology design in P1 to governance and feedback in P4.

 doi:10.20944/preprints202602.1970.v2

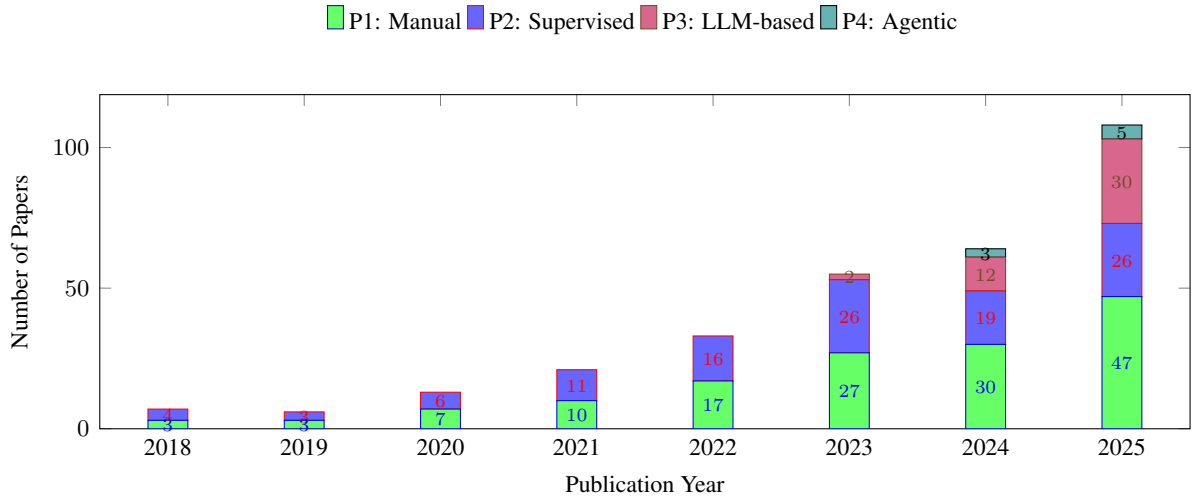

Figure A5. Dominant paradigms per year for the 310 KGs. The chart visualizes the steady baseline of manual/schema methods (P1) and the rapid rise of LLM-based methods (P3 & P4) from 2023 onwards.

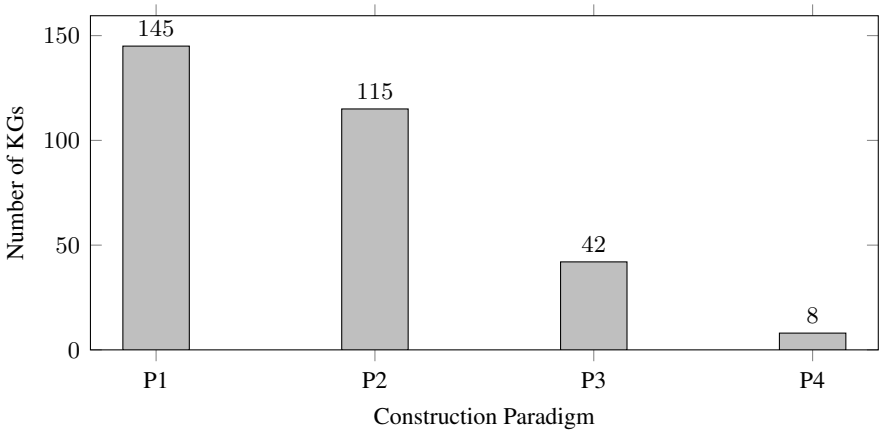

Figure A6. Overall distribution of dominant construction paradigms among the 310 KGs (2018–2025): **P1** (Manual/Ontology-first, 145 KGs; 46.8%) remains the most common for high-precision domains, followed closely by **P2** (Supervised/Rule-based NLP, 115 KGs; 37.1%). **P3** (LLM-based, 42 KGs; 13.5%) and **P4** (Agentic, 8 KGs; 2.6%) represent the emerging generation of resources.

Appendix F. KG Usage

Table A7 defines the categories and their typical interfaces to NLP pipelines (e.g., IE/EL and alignment for ingestion; KGQA/text-to-SPARQL/GraphRAG for querying; verification and provenance tracing for auditable outputs; and KG-enhanced representation learning for prediction). We use this taxonomy as an *orthogonal* axis to ESG topics: it is not meant to replace pillar–theme–focus labeling, but to clarify the functional role that the KG plays in end-to-end ESG applications.

Table A8–Table A11 summarize representative **KG→App** systems across the Environmental (E), Social (S), Governance (G), and Holistic (H) pillars, organized by the ESG-RFM pillar–theme–focus taxonomy. In addition to the topical placement, each row is annotated with the **KG usage taxonomy** that captures *how* the KG is operationalized in the downstream system (Appendix F, Table A7). We report: (i) the focus (taxonomy anchor), (ii) the dominant KG usage pattern(s) (INT/MON/QA/AUG/PRED/REC/PROV/DT/VIS/SHARE), (iii) the dominant NLP/ML blocks used to build/query the KG, and (iv) representative works.

 doi:10.20944/preprints202602.1970.v2

Table A7. A usage taxonomy for the KG→App stage: *how* a KG is operationalized inside ESG applications, and the corresponding NLP interface blocks.

Use Abbr	What the KG is doing	NLP interface (typical blocks)	Representative exemplars
INT	Unifies heterogeneous ESG evidence under a shared schema; enables cross-source joins and cross-standard comparability.	Entity linking/canonicalization; normalization (time/unit/boundary); schema-aligned retrieval and joins.	KnowWhereGraph; SustainGraph
MON	Tracks risks/events over time; supports alerts, screening, incident/claim monitoring, and compliance surveillance.	Streaming/periodic extraction; event linking; temporal KG updates; retrieval + classification.	ClimaFactsKG; FFHKG
QA	Answers compliance / reporting / evidence questions using KG structure as the retrieval spine (often with RAG).	Text-to-SPARQL/Cypher; KGQA; GraphRAG-style graph-conditioned retrieval and synthesis.	Glitter; GraphRAG-CEEDER
AUG	Feeds graph signals into ML models (KGE/GNN/HIN) for scoring, ranking, classification, and prediction.	IE/EL → KG; embeddings/GNN features; learning-to-rank or classification with graph signals.	ESG-HIN; OntoMetric
PRED	Uses graph structure for reasoning/prediction (rules, paths, link prediction, forecasting).	Rule/path reasoning; graph completion; causal/temporal linking; constrained inference.	Reactive CKG; GovGraph
REC	Matches entities/opportunities (suppliers, flows, interventions, curricula, materials) using KG constraints.	Candidate retrieval + constraint checking; KG similarity; sometimes LLM reasoning over KG context.	SmartSC; KG-based Skills
PROV	Supports auditability: lineage, traceability, and evidence chains for ESG claims/metrics.	Provenance capture; citation linking; qualified statements; validation and contradiction checks.	CarbonKG; TaxGraph
DT	Acts as a digital twin / world model for control, simulation, optimization, or dynamic planning.	Sensor/geo ingestion; state updates; tool execution; sometimes agentic querying/control.	TWA-dKG; WorldAvatar-dKG
VIS	Enables exploratory analytics, dashboards, and narrative inspection via graph views.	Graph summarization; community detection; lightweight IE for labels/edges.	ESG Narrative Networks
SHARE	Supports multi-stakeholder sharing and governed interoperability (data spaces, DPP ecosystems).	Schema alignment; access control metadata; provenance + versioning for exchange.	Solid-based ESG data spaces; DPP Ontologies

Table A8. Environmental (E) KG→App systems annotated with KG usage.

Theme	Focus	Use	NLP/ML backbone (build + query)	Representative systems
ClimChg	GHG, ProdFoot	PROV, INT, AUG	Ontology/ETL or IE over inventories; factor normalization; KGE/GNN for footprint completion; provenance-linked roll-ups	CF-KG; CarbonKG; E-Liability
	PhysClimRisk	MON, INT, PRED	NER/RE over news + geo-entity linking; GeoSPARQL; semantic-web geo KGs; dynamic/digital-twin updates	CC-KG; KnowWhereGraph; LinkClimate; TWA Flood
	ClimFin	AUG, INT	IE over reports; heterogeneous information networks; graph-based scoring (GNN/HIN); cross-source linkage	Green Premium KG; ESG-HIN
	ProdFoot	REC, INT, QA	Schema-guided retrieval of background flows; Cypher/SPARQL querying; LLM-assisted parameter completion	LCIKG; LCA-KG; Ecoinvent-KG
NatCap	BiodivLU	INT, MON, VIS	Scientific-text NER/EL; ontology alignment; geo-alignment to RS/GIS layers; hybrid RE + validation	Nature FIRST; GeoKG-HSA; FooDS
	RawSrc	PROV, INT, MON	Traceability ontologies; rule/path reasoning; supplier/entity linking; optional LLM IE	SCT-KG; SCN-KG; SCKG-MS

 doi:10.20944/preprints202602.1970.v2

Table A8. *Cont.*

Theme	Focus	Use	NLP/ML backbone (build + query)	Representative systems
	WaterSS	MON, INT, PRED	Semantic sensor/geo integration; event extraction; graph learning over spatiotemporal KGs	WHOW-KG; WQHR-KG; SmartWater-KG
PolWasCiru	EWaste	MON, PRED, INT	NER/RE from incident text; rule extraction; KGE/GNN for planning	EEE-GDKG; EVB-DKG; AIRQ-KG
	PackWaste	VIS, REC, INT	Topic/sentiment mining; ontology-driven metadata; stakeholder linkage	PRIME-KG; Circular Factory KG
SustSolTech	CleanTech	VIS, INT	Topic modeling; co-occurrence and citation graph analytics	Green-tech KG
	GreenBldg	DT, INT, PRED	Built-environment ontologies (Brick/BOT); rule reasoning; KGQA + optimization	Building KG; TWA-dKG; Housing KG
	ReEneg	DT, MON, INT	Multimodal KGC (SCADA/manuals); anomaly detection + KG reasoning	XAI4Wind; SCADA-KG; WorldAvatar-dKG

Table A9. Social (S) KG→App systems annotated with KG usage.

Theme	Focus	Use	NLP/ML backbone (build + query)	Representative systems
HumCap	OHS	MON, PRED, INT	Incident-report IE; causal/event linking; graph analytics; multimodal safety evidence	SCH-KG; Rail-Risk KG; FFHKG
	LearnDev	REC, INT, QA	NER/EL over course/job text; taxonomy alignment; KG recommenders	KG-based Skills
	WorkManRel	MON, AUG, PRED	Web/report IE; graph learning for attrition; link prediction for risk propagation	Industry-chain KG; Attrition KG; Labor-risk KG
ProdRespCustSafe	ChemSafe	MON, PRED, INT	Ontology-grounded IE; dictionary/rule mappings; KGE + rules	Soil-Toxic KG; Toxicity KG
	DataSec	QA, MON, PROV	Compliance-rule modeling; KG-augmented RAG; CTI entity/event extraction	PrivComp; AttacKG+; VulKG
	ProdQual	MON, QA, INT	Safety/quality IE; KG-backed reasoning; QA interfaces	HCP; QChsG
ComRigRisks	ComRel	INT, VIS	Ontology modeling; GeoSPARQL; web IE + graph analytics	People KG; Trafficking KG
	RiskSrc	MON, PROV, AUG	Event extraction; neurosymbolic reasoning; deep-tier link prediction	Deep-tier SC; Supplier-LLM KG
IncSolSocAcc	HealAcc	INT, AUG, QA	Tabular+text integration; KG-augmented querying; NL-to-Cypher agents	HealthEQKG; Health-Service KG
	NutriHeal	REC, PRED	KBQA over food/health ontologies; constraint-aware recommendation	FoodRec-KG

Table A10. Governance (G) KG→App systems annotated with KG usage.

Theme	Focus	Use	NLP/ML backbone (build + query)	Representative systems
CorpGov	OwnCtrl	PRED, MON, INT	Registry integration; rule/path reasoning (Datalog/Vadalog); temporal path analysis	VL-CO-KG; Reactive CKG; Equity KG
	FinRepQ	MON, AUG, PROV	Firm–statement graphs; heterogeneous GNNs; evidence tracing	CLFFG; AI-KG; SC-fraud KG
CorpCondInt	TaxTrans	QA, PROV, PRED	KG-centric retrieval; RAG-style explanation; audit signals	TaxGraph; TaxKG-RAG; Multimodal TaxKG
	EthicsAC	MON, PRED	Case IE; community detection; rule/pattern mining	DIS-KG; DIC-KG

 doi:10.20944/preprints202602.1970.v2

Table A11. Holistic (H) KG→App systems annotated with KG usage.

Theme	Focus	Use	NLP/ML backbone (build + query)	Representative systems
ESGIntAss	ESGRate	AUG, VIS, INT	Dynamic ESG networks; lexicon/sentiment KGs; graph-enhanced scoring	ESG Narrative Networks; ESG-HIN; ESGSenticNet
	ESGRept	QA, INT, PROV	Schema mapping across standards; RAG-style interpretation; provenance chains	ESRS–GRI RAG; Glitter; OntoMetric
	ESGData	QA, AUG	GraphRAG-style indexing; graph-conditioned retrieval + summarization	GraphRAG-CEEDER
SDFrame	SDGs	INT, QA, SHARE	Indicator/ontology alignment; KG-grounded retrieval across policy corpora	SustainGraph; ESGOnt; SDG-KG
SysSustMod	CircLoop	SHARE, DT, INT	Ontology networks for DPP ecosystems; governed data spaces; agentic tool chains	DPP Ontologies; Solid-based ESG data spaces; Agentic CE

Appendix G. Quantitative Trend Meta-Analysis

Based on the curated catalog of **337** KG-producing papers and the cross-engine query volume analysis defined by Table A5, we characterize how *research demand* and *knowledge-graph adoption* interact across the ESG Research Focus Map (ESG-RFM). Figure A7–Figure A10 present focus-level intensity and KG uptake across four major scholarly indices, while Figure A11–Figure A13 trace long-term temporal dynamics.

Three consistent trends emerge.

(T1) Demand–KG misalignment at the focus level. Across all engines, high-demand ESG foci—such as *Energy Consumption & Fuel Mix*, *Greenhouse Gas Emissions*, *Data Protection & Cybersecurity*, and *Digital Responsibility, Data & AI Ethics*—dominate absolute query volumes (Figure A7–Figure A9). However, KG adoption rates vary sharply across these foci, revealing systematic gaps where topic importance does not translate into structured knowledge representations.

(T2) Governance and data-centric topics exhibit higher KG penetration. While Environmental foci account for the largest absolute volumes, Governance- and data-heavy Social foci consistently show higher KG usage rates relative to their total volume, particularly in the ACL Anthology and Semantic Scholar (Figure A8, Figure A10). This pattern reflects the structural suitability of KGs for compliance reasoning, data governance, and formal policy modeling.

(T3) Accelerated KG uptake after 2021 across engines. Yearly trends reveal monotonic growth in ESG-related research since 2015, with a marked acceleration in KG-inclusive work after 2021 across Google Scholar, OpenAlex, and Semantic Scholar (Figure A11–Figure A13). Despite this acceleration, KG-based studies remain a small fraction of the overall ESG literature, underscoring substantial headroom for structured, machine-interpretable ESG knowledge.

Taken together, these quantitative signals position KG4ESG as a response to both *scale imbalance* (high-demand but weakly structured topics) and *methodological concentration* (KG-heavy adoption clustered in governance and analytics-oriented domains), motivating the need for a systematic atlas spanning all ESG pillars.

 doi:10.20944/preprints202602.1970.v2

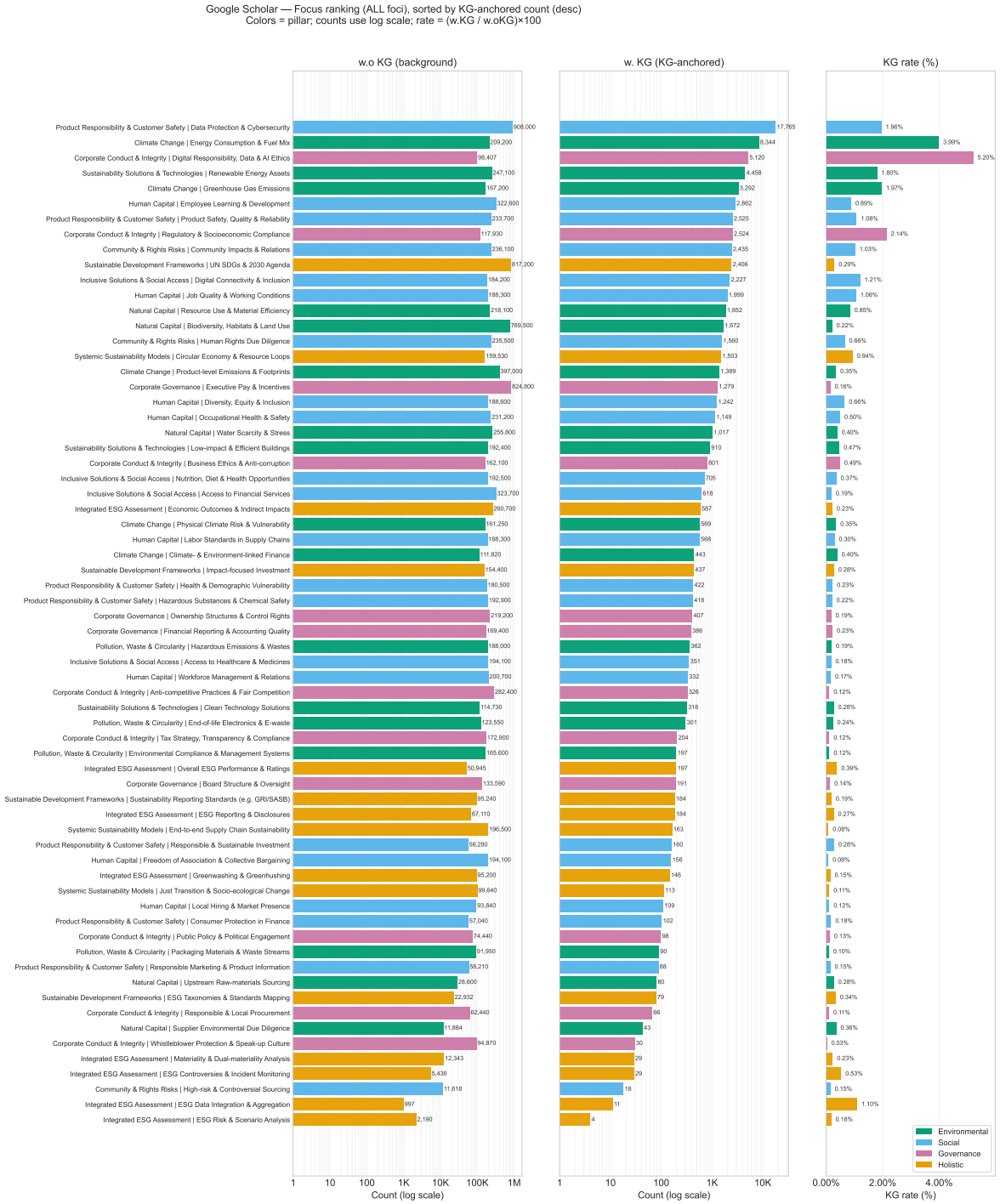

Figure A7. Query volume distribution across 66 ESG research focuses on Google Scholar. Bars indicate log-scaled absolute paper counts with and without knowledge graphs, while the overlaid curve reports the KG adoption rate. Focuses are ranked by total KG-inclusive volume.

 doi:10.20944/preprints202602.1970.v2

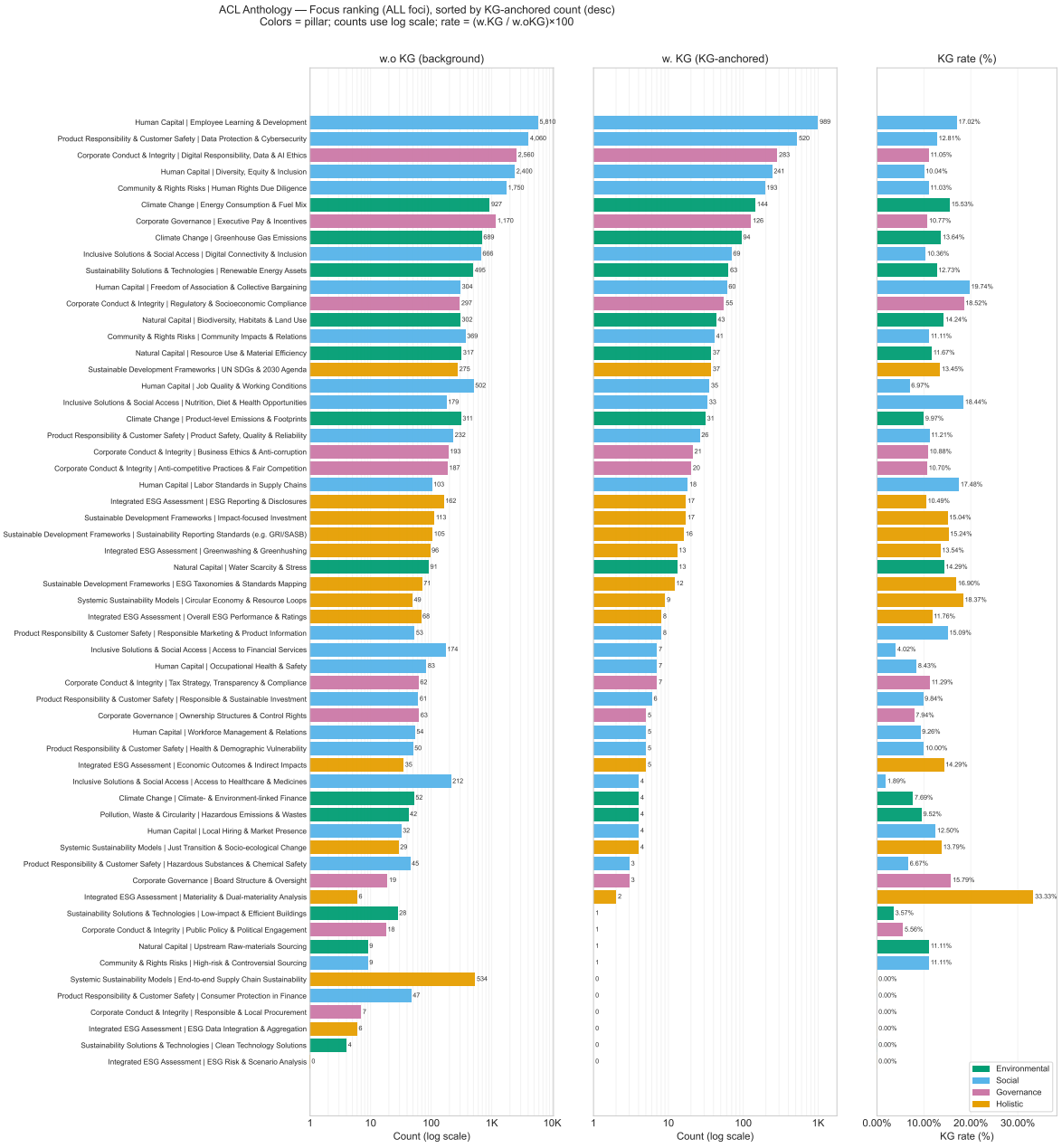

Figure A8. Query volume distribution across 66 ESG research focuses in the ACL Anthology. Log-scaled absolute counts contrast KG-based and non-KG works, with the KG usage rate overlaid and focuses sorted by total KG-inclusive volume.

 doi:10.20944/preprints202602.1970.v2

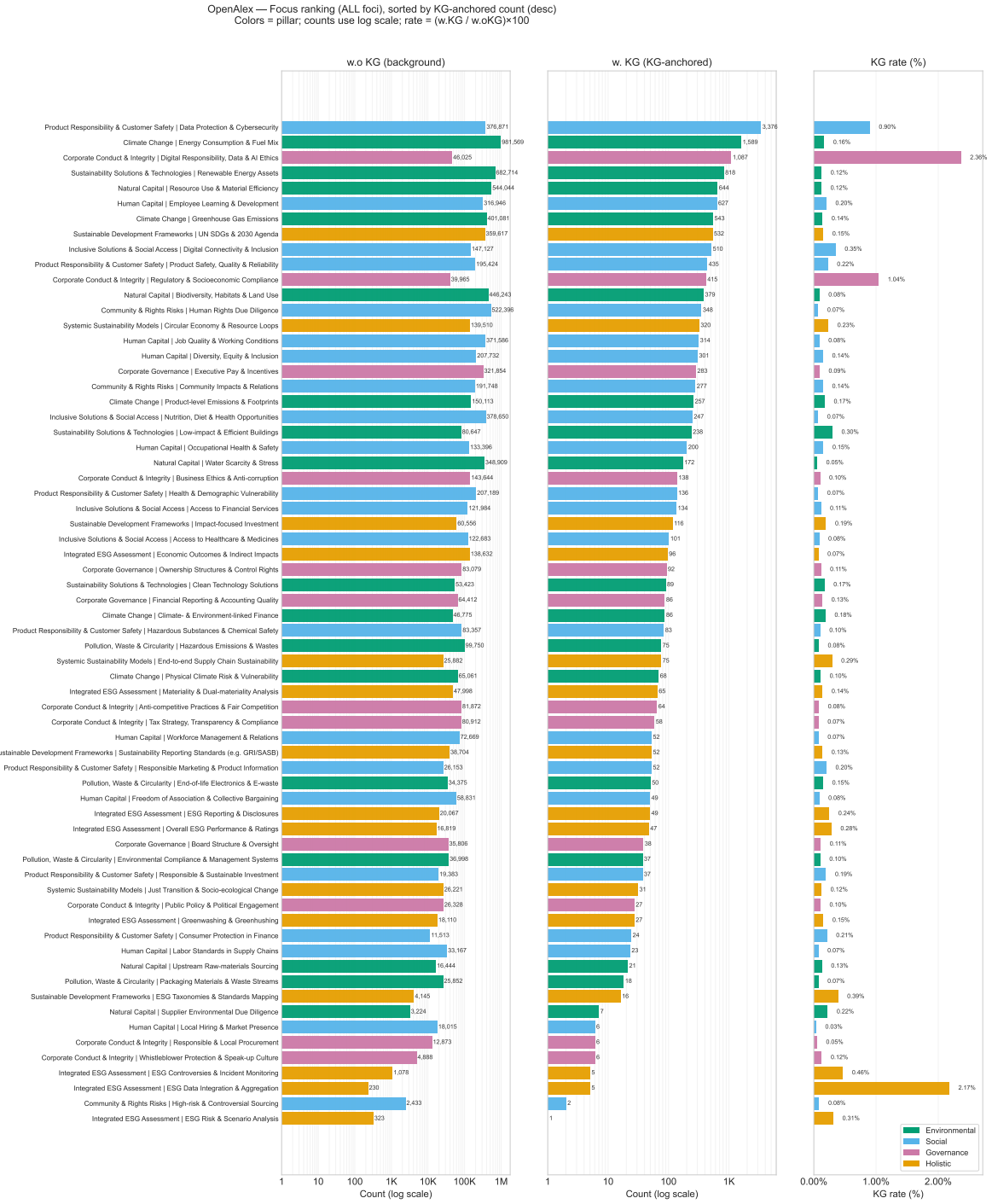

Figure A9. Query volume distribution across 66 ESG research focuses in OpenAlex. Absolute paper counts are shown on a log scale, together with the corresponding KG adoption rate, with focuses ranked by total KG-inclusive volume.

 doi:10.20944/preprints202602.1970.v2

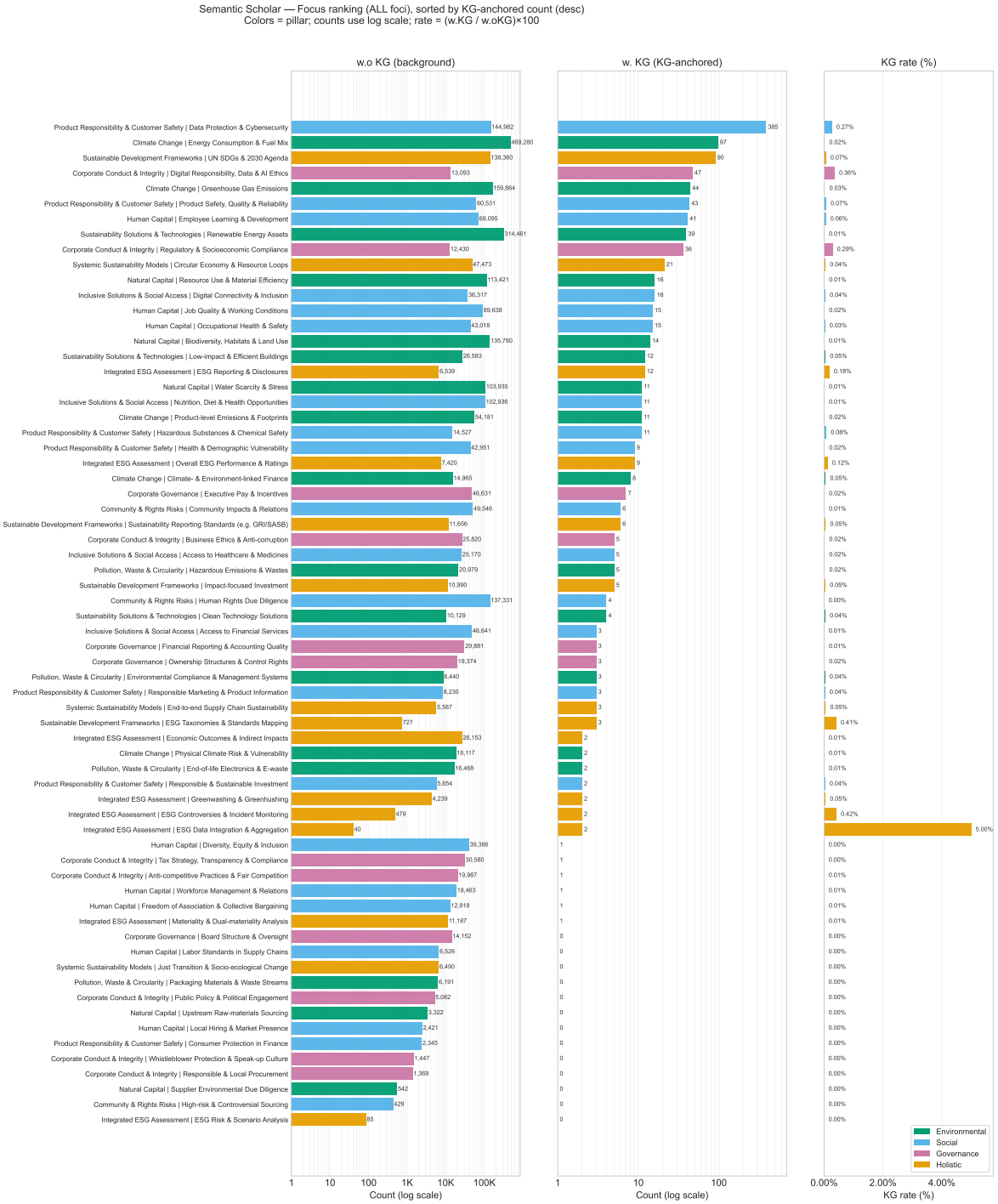

Figure A10. Query volume distribution across 66 ESG research focuses on Semantic Scholar. Log-scaled absolute volumes with and without knowledge graphs are shown, accompanied by KG usage rates, and focuses ranked by total KG-inclusive volume.

 doi:10.20944/preprints202602.1970.v2

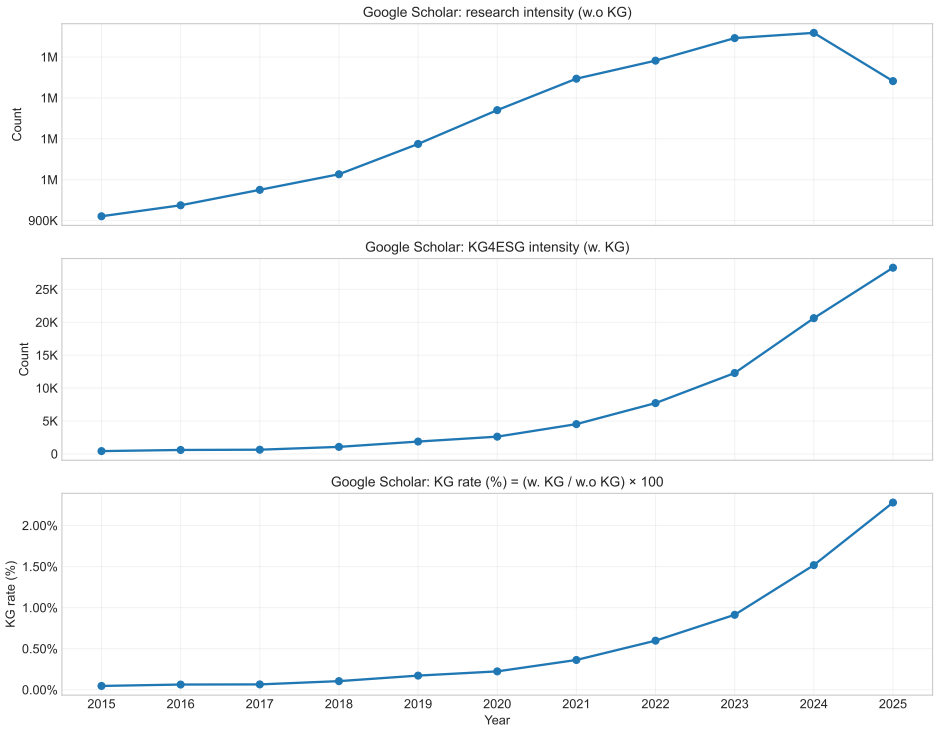

Figure A11. Yearly query volume trends for ESG-related research on Google Scholar. Log-scaled absolute counts of publications with and without knowledge graphs are shown by year, together with the corresponding KG adoption rate, revealing long-term growth dynamics and KG uptake over time.

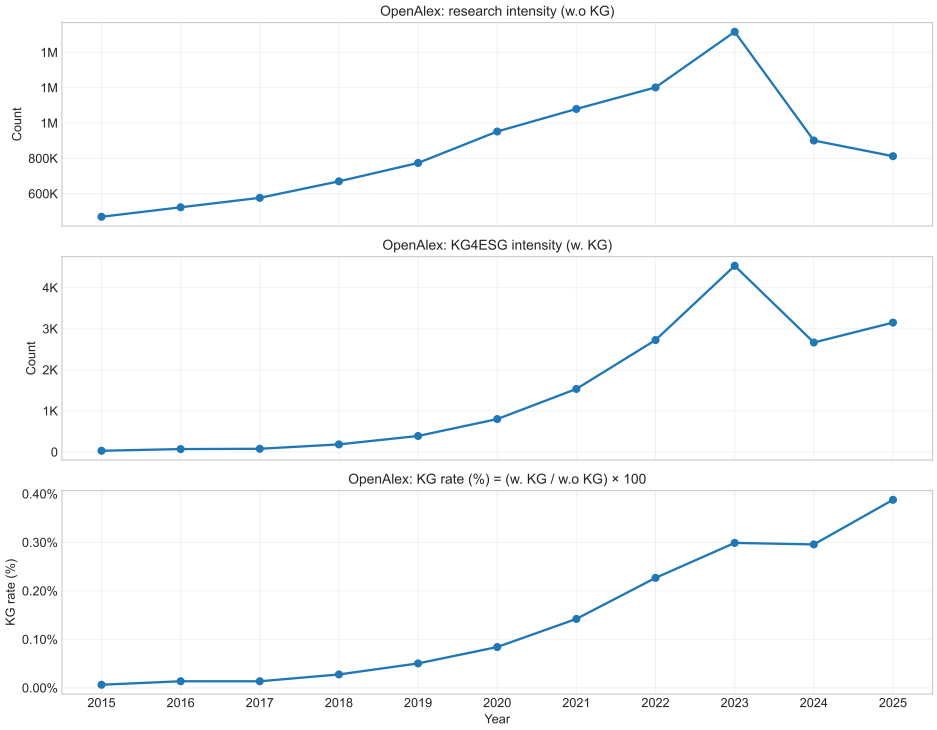

Figure A12. Yearly query volume trends for ESG-related research indexed in OpenAlex. Absolute publication counts with and without knowledge graphs are plotted on a log scale, alongside yearly KG usage rates, highlighting temporal shifts in corpus coverage and KG intensity.

 doi:10.20944/preprints202602.1970.v2

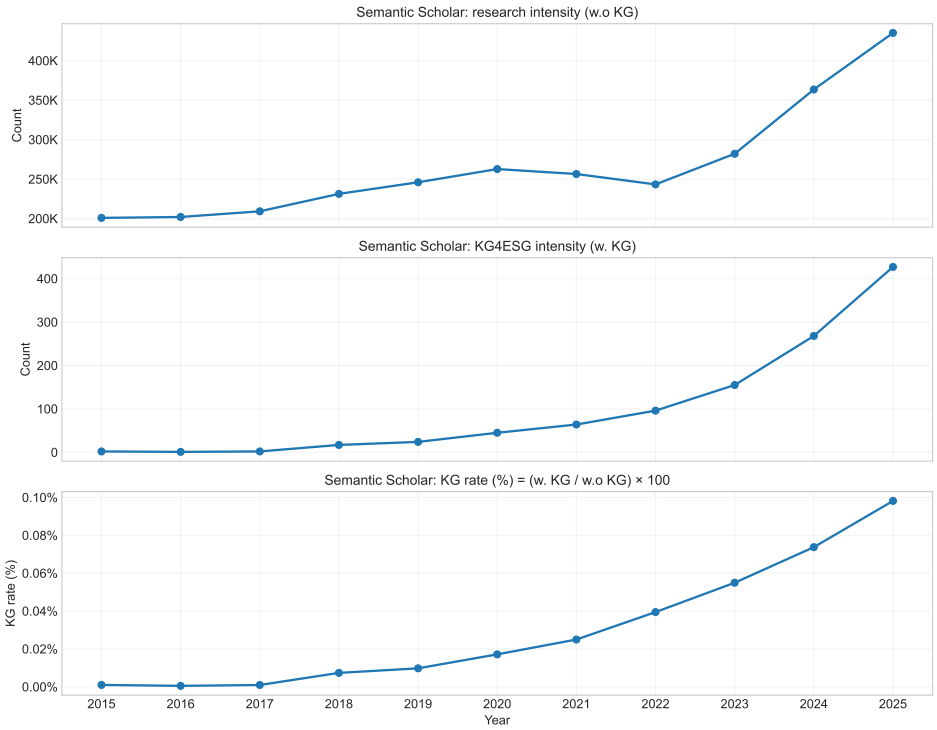

Figure A13. Yearly query volume trends for ESG-related research on Semantic Scholar. Log-scaled yearly counts contrast KG-based and non-KG works, with an overlaid KG adoption rate illustrating the evolution of KG usage across time.

Appendix H. Knowledge Graph Resource Catalog

This section catalogs **310 ESG-related knowledge graph (KG) resources** published between 2018 and 2025 (Table A12). Across the full survey, we review a total of **337 works**; the remaining **27 works** are *application-only studies* that leverage pre-existing KGs or external graph resources without constructing a new ESG-specific KG, and are therefore excluded from this construction-focused catalog.

The temporal distribution is strongly skewed toward recent years, highlighting the rapid acceleration of KG-based ESG research. Only **13 resources** appear before 2020, followed by sustained growth from 2020 onward (13 in 2020, 21 in 2021, 33 in 2022, and 55 in 2023), and a pronounced surge over the past two years, with **67 resources in 2024** and **108 in 2025 alone**. This pattern reflects both the maturation of enabling technologies (e.g., LLM-assisted extraction and graph tooling) and the rising demand for structured, traceable, and auditable ESG intelligence.

The catalog spans all four ESG pillars—Environmental (E), Social (S), Governance (G), and holistic/cross-cutting (H)—and covers a wide range of themes, including climate change and climate finance, natural capital and biodiversity, human capital and occupational health and safety, corporate governance and regulatory compliance, product responsibility (including cybersecurity and data security), sustainable solutions and technologies (e.g., renewable energy and green buildings), as well as supply-chain, circular-economy, and SDG-oriented frameworks.

Across the corpus, text and tabular sources dominate, but many resources integrate multiple modalities such as sensor and time-series streams, geospatial layers, imagery, graphs, CAD/BIM artifacts, and event logs. For each resource, the longtable reports ontology choices (existing, created, or mixed), construction paradigms, downstream KG usage and tasks, and scale; it also records—where stated—domain-expert involvement in ontology engineering, annotation, and validation. Finally, the availability field indicates the degree to which schemas, KG artifacts, extraction code, datasets, documentation, and access endpoints are publicly released, partially shared under restricted licenses, or remain proprietary and inaccessible to external researchers.

doi:10.20944/preprints202602.1970.v2

Table A12. Catalog of ESG-related knowledge-graph resources (2018–2025). Columns summarize ESG pillar (P), theme (T), focus (F), data sources and modalities, ontology, construction paradigm (Cstr; P1–P4), expert involvement, KG usage and tasks, graph scale, and availability.

Resource	Year	P	T	F	Data Source	Modality	Ontology	Cstr	Expert	KG Usage	Tasks	KG Scale	Availability
Wang et al. (2025)	2025	E	ClimChg	PhysClimRisk	10247 Academic Papers	Text	N.A.	P3	Yes (Val)	KGQA	RE; Align; Retr; QA; Other	Ent:23156; Rel:89472; Trip:127834	N.A.
Liu et al. (2025)	2025	S	ProdRespCustSafe	DataSec	63 Privacy Policies	Text	N.A.	P3	Yes (Annot,Val)	Bayesian Nets	Compliance Auditing	N.A.	N.A.
Sharma et al. (2025)	2025	S	ProdRespCustSafe	DataSec	Academic Papers	Text; Tabular	N.A.	P3	Yes (Corpus, Prompt, Sys Design)	RAG	Security, Privacy, Compliance	N.A.	N.A.
AEC-KG	2025	S	HumCap	OHS	6107 cases of China public accident reports and dispute cases	Text	Created	P2	Yes(Onto, Annot)	RiskMon	NER	N.A.	N.A.
Iwata et al. (2025)	2025	H	ESGIntAss	ESGIncid	800–2,000 News Articles	Text	Meta ESG Ontology	P3	Yes (Onto,Eval,Val)	Alignment	ESG controversy / violation detection; principle-level ESG classification (UNGC principles); interpretable ESG compliance analysis; downstream ESG trend tracking and risk prediction	N.A.	https://github.com/tsuyoshiiwataRR/NEWS_LLM_UT_RR
Zheng et al. (2025)	2025	E	PolWasCiru	HazWaste	Safety Management Data	Text; Tabular	Created	P1	Yes(Onto, Annot)	KGQA	Retr;QA	N.A.	N.A.
Yu et al. (2025)	2025	H	ESGIntAss	ESGRept	ESG regulatory PDFs	Text	ESGMKG	P3	Yes(Prompt, Val)	validated, provenance-preserving ESG regulatory representation	ESG compliance analysis, cross-framework alignment, RAG, sustainable-finance analytics, transparency tools	Document-level counts (42–68 entities, 42–65 relations per document)	N.A.
Wang et al. (2025)	2025	E	NatCap	WaterSS	64,512 risk inspection cases	Text; Geospatial; Event logs	Created	P2	Yes (Onto, Annot, Val)	mapping to risk evolution networks + spatio-temporal causal evolution analysis	causal propagation/diffusion analysis, temporal emergence rate, spatial associations, prevention/control strategy support	N.A.	N.A.
Felder et al. (2025)	2025	H	SysSustMod	SCSust	Enterprise Data(ERP / MES / CRM etc.)	Tabular; Geospatial; Graph; Other	a property graph ontology within the Neo4j GDB	P4	Yes(Onto,Prompt,Val)	multimodal route completion + optimization + emissions accounting for sustainable logistics/LCA	route enrichment, route optimization, emissions estimation, scenario benchmarking, sustainability reporting support	Ent:281; Trip:996	N.A.
ClimaFactsKG	2025	E	ClimChg	PhysClimRisk	SSkeptical Science (SkS) website articles, CimpleKG	Text; Graph	Schema.org, CARDS taxonomy	P1	Yes (Link)	Retr/QA	Cls; Retr; Other	N.A.	https://github.com/climatesense-project/climafacts-kg; https://purl.net/climatesense/climafactskg
DeBellis et al. (2025)	2025	H	ESGIntAss	GreenWH	Databases: 1.ASA & CAP advertising rulings; 2.Climate Change Litigation ; 3.DeSmog databases; 50 papers, web pages, book summaries, newspaper articles	Text; Tabular	Created and Dublin Core vocabulary; Gist Upper Model; Universal Moral Grammar ...	P1	Yes (Onto, Val)	RAG	Retr; QA; Other	N.A.	https://github.com/mdebellis/Climate_Obstruction
Huang and Song (2025)	2025	S	ProdRespCustSafe	ProdQual	Equipment maintenance documents, Operation manuals, Technical specifications, Maintenance records, Alarm logs	Text; Event logs	Created	P3	Yes(Onto, Prompt)	Semantic backbone	Fault diagnosis support	N.A.	
CPSKG	2025	S	ProdRespCustSafe	ChemSafe	Disaster Scenario Monitoring Data, Equipment Operation Data, Directories of Hazardous Chemicals, Accident Investigation Reports, Laws and Regulations, Emergency Plans	Text; Tabular; Sensor / time-series	CIPSR ontology	P3	Yes (Ont,Prompt, Val)	RAG	Emergency Decision Response	N.A.	N.A.

doi:10.20944/preprints202602.1970.v2

Table A12. *Cont.*

Resource	Year	P	T	F	Data Source	Modality	Ontology	Cstr	Expert	KG Usage	Tasks	KG Scale	Availability
CropDP-KG	2025	E	NatCap	BiodivLU	Chinese crop diseases and pests image-text database	Text; Tabular	Created	P2	Yes (Annot,Val)	Semantic backbone	QA; Decision Support	Ent:13840; Rel:7; Trip:21961	https://github.com/dadadaray/CropDP-KG/tree/Knowledge-System; https://doi.org/10.6084/m9.figshare.28015541
Cheng et al. (2025)	2025	S	ProdRespCustSafe	DataSec	150 real-world CTI reports	Text	MALOnt	P3	Yes (Annot, Prompt, Val)	Threat profiling / structured representation	CTI knowledge extraction / CSKG construction	N.A.	https://ctinexus.github.io/
Hou et al. (2025)	2025	E	PolWasCiru	EWaste	Solomon vehicle routing dataset	Tabular	A Schema	P1	Yes(Schema)	Knowledge memory for optimization	Dynamic vehicle routing optimization;Decision Support	N.A.	N.A.
Tabanao et al. (2025)	2025	E	NatCap	BiodivLU	COPIOUS biodiversity corpus; CABI Digital Library forestry compendium	Text	Not Mention	P2	Yes(Annot)	Reason/Pred	NER	N.A.	https://github.com/BiodivNER/BiodivNERModels
Wang et al. (2025)	2025	E	PolWasCiru	EWaste	Technical standards, policies, enterprise documents, disassembly reports; Historical databases of transformer models, materials, prices, and disassembly cases; Multimodal sensing (RGB-D images, gas concentration, resistance, oil quality, energy meters)	Text; Image; Sensor	A Schema	P3	Yes (Schema, Val)	KG+ML	NER; RE; Align; Retr; Other	Ent:3800; Trip:12500	N.A.
Fu et al. (2025)	2025	E	ClimChg	GHG	523 "dual carbon" policies; CNKI publications	Text	A policy ontology	P2	Yes (Onto, Annot,Val)	Reason/Pred	Retrival	Ent:23736; Trip:119744	https://github.com/yeahqiona/Dual-carbon-policies
Belova et al. (2025)	2025	H	SysSustMod	CircLoop	Enterprise data sources; Supplier-provided DPP data assets (KGs)	Tabular	Created	P1	Yes (Onto, Val)	Semantic backbone for DPP data assets	Carbon footprint tracking	N.A.	N.A.
Elderly Advantages Knowledge Graph	2025	S	ProdRespCustSafe	HealVuln	Elderly Interviews and Brainstorming Data	Text; Tabular	Created	P1	Yes (Ont,Val,Part)	Semantic backbone	Retrieval	N.A.	N.A.
EmeraldGraph	2025	H	ESGIntAss	GreenWH	37 publicly available ESG reports;KPI definitions; Claim collections	Text; Tabular; Image	Created ESG domain schema	P3	Yes(Onto, Val)	RAG	Greenwashing detection / claim verification	Ent:53748; Rel:59344	N.A.
Ji et al. (2025)	2025	E	ClimChg	GHG	Public research results / literature; Policy documents; Industry reports / technical reports	Text; Tabular	Created: ontology framework for carbon emission reduction	P2	Yes (Val)	multi-source integration and semantic organization layer	Carbon emission reduction pathway analysis and optimization; Decision Support	N.A.	N.A.
Energy Knowledge Graph (EKG)	2025	E	ClimChg	EnergyMix	Structured, semi-structured, and unstructured manufacturing data; expert knowledge; third-party knowledge base	Text; Tabular; Sensor / time-series	manufacturing energy management ontology via Stanford Seven-Step Method	P2	Yes (Onto, Val)	Semantic backbone	QA; Retr; Rec; Decision Support	N.A.	N.A.
Yang et al. (2025)	2025	E	PolWasCiru	EnvComp	Policy corpus for Niger	Text	Created: Enhanced ontology	P3	Yes (Annot,Val)	KG+ML	NER; RE; QA; Other	Ent:61912; Rel:81389	N.A.
Equity Knowledge Graph (EKG)	2025	G	CorpGov	OwnCtrl	CSMAR database;Online news;Official website	Text; Tabular	Created	P1	Yes(Data)	Unified representation; Learning substrate; Explainability	Actual Controller (AC) disclosure framed as control relation prediction between a shareholder and a company	Ent:1262542; Rel:1869448	N.A.
ESGSenticNet	2025	H	ESGIntAss	ESGRept	1,998 SGX company sustainability reports	Text	Created	P3	Yes (Ont,Annot,Val)	knowledge base	ESG topic analysis	Ent:23245; Trip:44232	N.A.
PREJUST4WOMAN	2025	S	ComRigRisks	HRDD	73 ECHR judgments/decisions	Text; Tabular; Graph	Created	P3	Yes (Annot,Val)	Knowledge-intensive legal information access	SPARQL-based legal querying; support legal decision-making; potential predictive-justice link prediction over KG	Ent:5185; Rel:22; Trip:10325	https://github.com/Fra3005/PreJust4Womans; https://github.com/Fra3005/PreJust4Womans/blob/main/Commons/final_onto.ttl;https://lod-cloud.net/dataset/PREJUST4WOMAN_PROJECT

doi:10.20944/preprints202602.1970.v2

Table A12. *Cont.*

Resource	Year	P	T	F	Data Source	Modality	Ontology	Cstr	Expert	KG Usage	Tasks	KG Scale	Availability
ESKG	2025	S	ProdRespCust Safe	ProdQual	Supply-chain node I-O data; ERP/MES; quality/inspection/after-sales orders; unstructured text; multi-modal data	Text; Tabular; Sensor / time-series; Image / raster; Event logs; Other	Created	P1	Yes(Onto)	Integration, reasoning, and feedback control for full-chain product quality tracing	Point/chain/network tracing and quality feedback control	N.A.	N.A.
FFHKG	2025	S	HumCap	OHS	1,097 accident reports	Text	AcciMap-based FFHKG ontology	P3	Yes (Annot,Val)	Causal knowledge representation	Automated risk factor extraction, causal analysis, and safety risk assessment	Ent:14466; Trip:49788	N.A.
Financial Knowledge Graph (FKG)	2025	G	CorpGov	FinRepQ	CSMAR databases	Tabular; Graph	A schema	P1	Yes(Schema)	relational substrate	Cls; GraphComp	MBM (Ent:790374; Trip:5580916);SME:(Ent:295461; Trip:2311465); GEM(Ent:220060; Trip:1651936)	https://github.com/DongskyGit/KeHGN-R
Firmographica	2025	E	NatCap	BiodivLU	FINRA short-interest reports; SEC 13-D/13-G filings and insider trades; LSEG PermID; Yahoo Finance indicators (216 banking firms, Jan–Oct 2024)	Tabular	A schema	P1	Unk	Ownership network representation and feature extraction for risk modeling	Short-selling position and volatility prediction	N.A.	N.A.
Wu et al. (2025)	2025	S	IncSolSocAcc	NutriHeal	Digital receipts; Food Composition Database (FCD); Product metadata; Nutritional standards; User survey data	Tabular	NutriLink ontology	P1	Yes (Ont)	Semantic integration and querying backbone for automated nutrition assessment and recommendations	Fully automated diet counseling, nutrition assessment, and personalized dietary recommendations	N.A.	https://purl.org/nutrilink;11https://purl.org/foodcoach-system
Gao et al. (2025)	2025	S	IncSolSocAcc	NutriHeal	12,500 recipes from two professional nutritional cookbooks; nutrients, ingredients, dietary constraints; multilingual expansion	Text	A schema	P3	Yes(Onto	Semantic, explainable knowledge backbone for LLM-grounded dietary guidance	Personalized recipe recommendation, constraint-aware dietary guidance, interactive nutrition exploration	https://github.com/astridesa/HealthGenie	
FooDS	2025	E	NatCap	BiodivLU	Soil, vegetation, GPS collar sensor data, and camera trap image metadata from wildlife observatories.	Tabular; Sensor / time-series; Image / raster; Geospatial	Forest Observatory Ontology (FOO)	P1	Yes (Ont,Val)	Semantic integration, reasoning, and linked-data access for wildlife monitoring	Wildlife tracking, habitat analysis, hazard detection, and conservation decision support	4 KGs (soil, vegetation, GPS collar, camera trap)	https://w3id.org/def/foo; https://w3id.org/def/fooDocs https://w3id.org/def/fooDS; andhttps://ontology.forest-observatory.org
ForestFoodKG-KG	2025	E	ClimChg	ProdFoot	Lit: Web-scraped documents / web pages; Scientific literature / scholarly articles	Text; Tabular	Existing: PROV	P2	Yes (Annot,Val)	Integration	NER; RE	Ent:4492; Trip:14130	https://github.com/dadadaray/FTAND
Zhao et al. (2025)	2025	H	SysSustMod	CircLoop	Ecoinvent Waste Treatment & Recycling (3,896 entries), multi-code systems (EWC/HS/NACE/CPA/ISIC/SSIC/WZ/CPC), GWP100 factors, facility/provider/receiver metadata, regulatory constraints	Text; Tabular	Created: incorporates ontology	P1	Yes (Onto, Annot, Val)	Fact-checked grounding and multi-hop reasoning for circular economy waste-to-resource decision support	Single-hop QA, multi-hop synergy discovery, GWP100 numeric retrieval and ranking, regulatory-ready planning support	Ent:117380; Trip:753145	N.A.
Spatiotemporal Knowledge Graph (STKG)	2025	H	ESGIntAss	EconImp	AIS trajectories; berth/port administrative records; vessel attributes; archive text; remote sensing imagery	Text; Tabular; Sensor / time-series; Image / raster; Geospatial	Created	P1	Yes(Onto)	Structured representation and reasoning backbone for berth allocation and optimization	Intelligent berth recommendation, utilization optimization, waiting-time reduction, port decision support	N.A.	N.A.
GeoKG (Habitat GeoKG)	2025	E	NatCap	BiodivLU	Species occurrence records; remote sensing (NDVI, land cover, humidity); DEM-derived topography; hydrology; road networks	Tabular; Sensor / time-series; Image / raster; Geospatial	Wetland Monitoring Ontology (WMO)	P2	Unk	Semantic integration and feature generation to enhance habitat suitability modeling	Habitat suitability prediction, spatial generalization, and environmental factor analysis	N.A.	N.A.

doi:10.20944/preprints202602.1970.v2

Table A12. *Cont.*

Resource	Year	P	T	F	Data Source	Modality	Ontology	Cstr	Expert	KG Usage	Tasks	KG Scale	Availability
GeoOutageKG	2025	H	SysSustMod	JustTrans	NASA Black Marble NTL imagery (county-masked), EAGLE-I county outage time series (15-min), derived outage severity maps, plus DBpedia/GEOSatDB/media ontologies for linking	Tabular; Sensor / time-series; Image / raster; Geospatial; Graph	GeoOutageOnto	P1	Yes(Onto)	Multimodal geospatiotemporal integration and multi-resolution reasoning for outage analysis	Outage severity evaluation, vulnerability/disparity analysis, and grid decision support	10,965,241 instances; 88,971,709 RDF triples; OutageRecord 10,635,995; NTLImage 313,702; OutageMap 15,544	https://doi.org/10.17605/OSF.IO/QVD8B; https://purl.org/geooutagekg; https://purl.org/geooutageonto
GovGraph	2025	G	CorpCondInt	RegComp	385 social governance innovation cases (29 regions), think tank expert opinions, 1,390 CNKI academic papers	Text	Created	P3	Yes(Onto, Annot, Val)	Structured representation and inference for forecasting social governance innovation networks	Entity extraction, node prediction, link prediction, and governance pathway discovery	Extracted entities: 14,386 stakeholders; 18,318 objects; 18,852 means; 5,624 environment; predicted 352 new nodes and 25,098 links (model outputs)	N.A.
Ushio et al. (2025)	2025	E	ClimChg	ClimFin	Administrative meeting minutes	Text; Tabular	Not named	P1	Yes(Onto)	Semantic grounding and graph-based context expansion for QA	Administrative QA, relational exploration, aggregation, and policy analysis	N.A.	N.A.
Hazardous Chemical Accident Knowledge Graph (HCAKG)	2025	S	ProdRespCustSafe	ChemSafe	461 unstructured HCA investigation reports; safety regulations and chemical standards	Text; Tabular	HCA Ontology	P2	Yes (Ont,Val)	Structured risk knowledge representation and quantitative accident risk analysis	Risk factor extraction, causal network analysis, and safety decision support	N.A.	N.A.
HealthEQKG	2025	S	IncSolSocAcc	HealAcc	CMS Provider Data Catalog, Area Deprivation Index, USPS Crosswalk, DBpedia	Text; Tabular; Graph	HealthEQKG Ontology	P1	Yes(Onto,Annot,Val)	Semantic, queryable foundation for national-scale health equity research	Health equity analysis, workforce disparity assessment, and policy-support querying	198,997 nodes; 318,978 relationships; 72,658 physicians; 28,346 ADI-linked ZIP areas	https://doi.org/10.5281/zenodo.15708535; http://healtheqkg.myftp.org/sparql; https://github.com/navapatn/HealthEQKG
KnowWhereGraph (KWG)	2025	E	PolWasCiru	EnvComp	30+ open government and third-party geospatial datasets across environment, hazards, public health, transportation, and humanitarian relief	Text; Tabular; Image / raster; Geospatial; Graph	KnowWhereGraph Ontology	P1	Yes (Ont,Part)	Large-scale geo-semantic integration, enrichment, and reasoning	Disaster response, supply-chain analysis, land valuation, expert discovery, and geo-enrichment	Trip:29B	https://github.com/KnowWhereGraph; https://knowwheregraph.org/; https://stko-kwg.geog.ucsb.edu/graphdb/
Person–Job Temporal Knowledge Graph	2025	S	HumCap	LearnDev	2000 unstructured Chinese resumes; external KBs for alignment	Text	A Schema	P3	Yes(Schema, Val)	Temporal modeling of person–job dynamics and reasoning	Career analysis, relational reasoning, and person–job matching	N.A.	https://tianchi.aliyun.com/competition/entrance/231771/information; https://github.com/spohon/PJLLMsTKG-
JobEdKG and T-JobEdKG	2025	S	HumCap	WorkManRel	99,676 job ads (Rekrute.com), 10,180 MOOCs (Coursera), ESCO/ROME skill taxonomies	Text; Tabular	A custom metamodel	P2	Yes(Onto)	Temporal representation and completion of job–skill relations	Skill demand forecasting and temporal link prediction	Ent:55718; Rel:12; Trip:1296374	Sch:partial; KG:yes; Data:yes; Code:yes; EP:3; https://github.com/team611/JobEd; https://github.com/BahajAdil/TempTorchKGE; https://shorturl.at/divxC
ECOLOPES KG	2025	E	NatCap	BiodivLU	GLoBI, GBIF (API + backbone RDF), Wikidata, Nature FIRST KG; local Vienna datasets; internal PFG/AFG; voxel model in Postgres; CAD/Grasshopper node networks	Text; Tabular; Geospatial; Graph; CAD/BIM	EIM (ECOLOPES Information Model) ontology	P1	Yes (Ont,Part)	Semantic mediator and constraint-validation substrate linking ecology data with architectural geometry for biodiversity-informed design	Solar/proximity/prey constraint validation and biodiversity-aware building/site design decision support	Dataset-level triple counts reported (e.g., 2,454,463 materialized; 136,865,946 federated; 1,329,096 virtualized)	https://github.com/aahmeti/Ecolopes

doi:10.20944/preprints202602.1970.v2

Table A12. *Cont.*

Resource	Year	P	T	F	Data Source	Modality	Ontology	Cstr	Expert	KG Usage	Tasks	KG Scale	Availability
Regional Graph (RG) + Intent Graph (IG)	2025	H	SDFrame	SDGs	County-level indicators (geographic, economic, environmental, cultural), 94 SDPs, 7,830 historical region–SDP interactions	Text; Tabular; Geospatial	Not Mention	P1	Unk	Interpretable semantic backbone for regional SDP recommendation	SDP recommendation, explainable policy pathway discovery	2,596 regions; 94 SDPs; 7,830 interactions; RG attributes reduced from 39,744 to 1,669 after pruning	KG:partial; Data:partial
Benjira et al. (2025)	2025	H	SDFrame	SDGs	UN SDG metadata text; INSEE and French open datasets (census, facilities); Sport Ministry facilities; other open data sources used in examples	Text; Tabular; Geospatial	A SDG schema	P3	Yes (Ont,Val)	Structured schema alignment and indicator computation via queryable mappings	LLM-augmented schema mapping and SDG indicator computation	Case-level excerpt for Goal 11: 508 nodes	N.A.
IoT-Reg Knowledge Graph	2025	S	ProdRespCust Safe	DataSec	Synthetic IoT GDPR scenarios; GDPR articles encoded as deontic rules	Text; Other	IoT-Reg Ontology	P1	Yes (Onto, Val)	Structured, deontic grounding of IoT GDPR compliance reasoning	Natural-language compliance QA, automated compliance checking, explainable regulatory advice	N.A.	N.A.
Hassan Nassar et al. (2025)	2025	H	ESGIntAss	ESGData	1,098 ESG news articles from News-API/Guardian/BBC/CNN; 550-company registry (FTSE 100/250, ASX 200)	Text; Tabular	Not named	P2	Yes(Schema)	Semantic integration and inferential ESG intelligence from real-time news	ESG sentiment analysis, sector/index comparison, explainable ESG insights	Ent:3942; Rel:6; Trip:7007	https://github.com/WCKDNaz/KG-ESG-UEL
Papageorgiou et al. (2025)	2025	E	NatCap	WaterSS	Literature-derived process/model knowledge; PMF XML files from simulation tools; CSV KPI datasets; operational sensor/ERP data (supported)	Text; Tabular; Sensor / time-series; Geospatial; Other	Not named	P1	Yes (Ont)	Integrating processes + models + KPIs for explainable knowledge management and query answering	Process-related QA, model/KPI discovery, and process optimization support	N.A.	N.A.
knowledge graph structure	2025	E	SustSolTech	ReEneg	Forecasted wind/solar output, hydro inflow, and power load (SWLSTM-GPR + Monte Carlo); system constraints from Yalong River WSH base	Tabular; Sensor / time-series	Created: knowledge graph schema	P1	Yes(Onto)	Interpretable, traversable dispatch-rule representation and optimization container	Multi-objective WSH dispatch rule optimization and load-matching decision support	N.A.	https://github.com/zzdzzdzzdzzd/KG-for-Yalong-River-Operation-Rule
Lin and Liao (2025)	2025	H	SDFrame	SDGs	UN SDG indicator docs; Elsevier SDG keywords; TED transcripts (269 pilot; 1,127 formal)	Text; Tabular	Not named	P3	Yes(Curate,Prompt)	Speculative SDG interconnection analysis and new-goal ideation	Speculative SDG interconnection analysis and new-goal ideation	SDG relationship mining, simulated discussion analysis, new SDG goal generation, visualization	https://kg-web-4-0.vercel.app/about.html
LCA Knowledge Graph	2025	E	ClimChg	ProdFoot	Ecoinvent datasets and glossary; enterprise internal data sources mapped to indicators	Tabular; Graph	Ecoinvent vocabulary / Data Glossary (JSON-LD), KPIOnto	P1	Yes	Formalizing and computing LCA impact indicators and enabling consistent comparisons across organizations	Indicator computation, formula derivation, missing-data identification, and multi-enterprise dependency analysis	N.A.	https://glossary.ecoinvent.org/
LCAIM Knowledge Graph	2025	G	CorpGov	OwnCtrl	Asset registers, BMS energy data, supplier databases, analysis metadata	Tabular; Sensor / time-series	LCAIM Ontology	P1	Yes (Ont,Val,Part,Reg)	Integrated life-cycle analysis, reporting, enrichment, and auditing	LCC analysis, sustainability reporting, asset portfolio decision support, data quality checks	N.A.	Link Not Found
Dolha et al. (2025)	2025	H	ESGIntAss	ESGRept	ESG-enriched BPMN process models; ESG regulations/standards as semantic drivers	Text; Graph; Other	Existing: RDF	P1	Yes (Part)	Process-centric ESG knowledge management and LLM grounding	ESG analysis, reporting, decision support, and process redesign	N.A.	N.A.
Cai et al. (2025)	2025	H	ESGIntAss	ESGRate	1,500 Fortune Global 500 sustainability reports (2020–2023); standards-based KGs; 300-report ground truth sample	Text; Tabular	Created: and ontology	P4	Yes (Annot,Val)	Semantic alignment of sustainability metrics to support benchmark-ready extraction	Benchmark dataset creation, extraction, data-quality scoring, and rating-method validation	N.A.	N.A.
Aivalis et al. (2025)	2025	E	ClimChg	PhysClimRisk	ReliefWeb flood reports + metadata; OSM natural features; Microsoft Planetary Computer RGB/NIR/NDWI satellite imagery	Text; Tabular; Image / raster; Geospatial	Not named	P3	Yes()	Transparent, queryable disaster representations enabling interpretable comparison and decision support	Disaster impact querying, event similarity retrieval, geospatial and image-based flood analysis	N.A.	https://github.com/teoaivalis/XtremeKG

doi:10.20944/preprints202602.1970.v2

Table A12. *Cont.*

Resource	Year	P	T	F	Data Source	Modality	Ontology	Cstr	Expert	KG Usage	Tasks	KG Scale	Availability
Multi-modal Process Knowledge Graph for Wind Turbines (MPKG-WT) Thapa et al. (2025)	2025	E	SustSolTech	ReEneg	Industrial wind-turbine manuals, images, CAD models, JSON process documents	Text; Tabular; Image / raster; Graph; CAD/BIM	MPKG-WT Ontology	P2	Yes (Ont,Annot,Val)	Semantic integration and grounding for assembly process QA	Multi-modal QA, assembly process design support, and knowledge reuse	Ent:1576; Trip:3480	N.A.
Zhang (2025)	2025	G	CorpCondInt	TaxTrans	Subproject research outputs; files/metadata in Kadi4Mat; structured RDF in triple store	Tabular; Other	Not named	P1	Yes (Ont,Part)	FAIR-aligned, interoperable knowledge exchange and future automated circular-factory operations	Integration, querying, metadata retrieval, and automation support (planned)	N.A.	N.A.
METRIN-KG	2025	E	NatCap	BiodivLU	Invoice images, ERP texts/emails/policies, transactional tables	Text; Tabular; Image / raster	Not mentioned	P2	Yes	Semantic integration and reasoning for intelligent tax compliance	Compliance risk detection, audit automation, explainable tax governance	N.A.	N.A.
Multimodal Knowledge Graph (MMKG)	2025	S	ProdRespCust Safe	ChemSafe	ENPKG metabolomes; TRY plant traits; GloBI biotic interactions; Wikidata taxonomy	Tabular	Earth Metabolome Ontology (EMI ontology)	P1	Yes(Onto)	Integrated, FAIR, queryable biodiver-sity–metabolome–interaction knowledge	Ecological analysis, natural-product discovery, conservation and agriculture research	1.8M traits, 12.9M interactions, 65k species mapped; ENPKG 1,600 extracts	https://github.com/earth-metabolome-initiative/metrin-kg; https://kg.earthmetabolome.org/metrin/; https://github.com/earth-metabolome-initiative/earth_metabolome_ontology; https://doi.org/10.5281/zenodo.16874620
Cheng et al. (2025)	2025	E	PolWasCiru	HazWaste	Chinese MSDS PDFs; molecular structure images; ChEMBL images (pre-training)	Text; Image / raster	A Domain Schema	P2	Yes	Structural knowledge injection to guide vision–language alignment	Image-to-text retrieval for hazardous chemical identification; emergency safety support	31,264 nodes; 108,823 relations; 2,380 chemical entities	N.A.
Yin et al. (2025)	2025	E	PolWasCiru	EWaste	Inventory of Hazardous Chemicals (2015) + GHS (CAS mapping)	Text; Tabular	Created: OntoRXN ontology	P2	Yes	Domain-knowledge embedding to improve hazardous chemical text recognition and attribute classification	Recovery identification and hazard-category classification (21 subcategories)	Trip:18481	N.A.
Housing Passport Knowledge Graph (HPKG)	2025	H	SysSustMod	CircLoop	Battery disassembly manuals; experimental disassembly measurements; derived cost/carbon attributes	Text; Tabular; Event logs	Not Mention	P1	Yes	Representing and reasoning over disassembly dependencies for optimization	Disassembly sequence optimization for power-battery remanufacturing	N.A.	N.A.
Sheng et al. (2025)	2025	E	ClimChg	GHG	BIM archetypes (Revit), GIS geospatial + socioeconomic datasets (ArcGIS), multi-provider Canadian open data; SDG-linked context	Tabular; Geospatial; CAD/BIM	Data Homebase ontology	P1	Yes (Ont)	Standardized semantic infrastructure enabling multi-scale housing circularity and affordability analysis	Housing passport generation, spatiotemporal analytics, footprint assessment, circular material scenario evaluation	N.A.	N.A.
Yang et al. (2025)	2025	E	NatCap	WaterSS	Space-air-ground-social multi-source sensing; structured carbon flux/CO_2/trajectory data; unstructured policy/reports/literature	Text; Tabular; Sensor / time-series; Geospatial	Not Named	P3	Yes	Data-driven carbon governance with dynamic monitoring and fine-grained analysis	Sequestration assessment, emissions trading support, spatiotemporal carbon analytics	1M nodes; 1.3M edges	N.A.
NatureKG	2025	E	ClimChg	ClimFin	Inspection images/text from Middle Route Project; risk manuals/guidelines; Milvus/web sources for MAAR retrieval	Text; Image / raster	Not Named	P2	Yes (Annot)	Grounding and structured retrieval for risk decision-making generation	Multimodal risk identification and contingency-plan generation via MAAR	Ent:14228	N.A.
AlMahri et al. (2025)	2025	H	SysSustMod	SCSust	ENCORE, SBTN, academic/gray literature (built environment focus)	Text; Tabular	Not named	P1	Yes (Ont,Val)	Structured representation of nature-finance dependen-cies/risks/actions and LLM-grounded querying	Text2Cypher QA over NatureKG; evidence-backed nature-finance insights	Ent:320; Trip:540	https://zenodo.org/records/16965298
Heus et al. (2025)	2025	H	SysSustMod	SCSust	Public web text (news, websites, Wikipedia)	Text	Not mentioned	P3	Yes	Enhancing multi-tier supply chain visibility without direct partner information sharing	Supply chain mapping, risk analysis, alternative sourcing identification	Ent:1158; Rel:641	N.A.
					MSCI MAC factors + methodology; LexisNexis news; synthetic supply chain KG; FAISS indices	Text; Tabular; Graph	Not Mentioned	P4	Unk	Centrality-guided path discovery and explainable prompt scaffolding for risk narratives	Agentic supply chain risk analysis with multi-modal evidence synthesis	N.A.	N.A.

doi:10.20944/preprints202602.1970.v2

Table A12. *Cont.*

Resource	Year	P	T	F	Data Source	Modality	Ontology	Cstr	Expert	KG Usage	Tasks	KG Scale	Availability
Badmus et al. (2025)	2025	G	CorpGov	FinRepQ	Multilingual financial reports, XBRL filings, MD&A, footnotes, regulatory standards	Text; Tabular; Image / raster	Both: master ontology; shared ontology	P2	Yes (Ont,Val)	Semantic harmonization, compliance reasoning, explainable financial governance	Cross-border reporting harmonization, summarization, compliance checking, risk and investor analysis	Trip:10000000	N.A.
Ashworth et al. (2025)	2025	E	NatCap	BiodivLU	Forage Data Hub (52,997 entries, 108 locations, 51 years), weather/soil data, EPA ecoregions	Tabular	Not Mentioned	P1	Yes(Onto)	Integrated semantic-spatial reasoning on forage system resilience	Climate-resilient forage system identification and regional decision support	N.A.	https://doi.org/10.15482/USDA.ADC/1529174
Cotti et al. (2025)	2025	S	ProdRespCustSafe	DataSec	AIT-LDS benchmark logs; Cowrie honeypot logs (Aug 4–14, 2025); optional context metadata	Text; Event logs	Not Mentioned	P4	Yes (Annot)	Ontology-grounded CTI extraction and explainable evidence for ATT&CK mapping	Log-to-KG CTI extraction and MITRE ATT&CK tactics prediction	N.A.	N.A.
Zheng et al. (2025)	2025	H	SysSustMod	SCSust	Open data (news/government reports); case study: >7,000 cyclone-related news articles via open API	Text	NaturalHazard, NaturalHazardSupplyChainImpact, Conflict/MilitaryConflict; user-specified ontology	P3	Yes	Structured semantic risk monitoring and explainable impact annotation	Risk extraction and impact-aware risk alerting for procurement	Ent:30881; Trip:51893	N.A. 1
Qiao et al. (2025)	2025	S	ProdRespCustSafe	DataSec	Other: Unspecified data source (see data_sources_summary)	Text; Tabular; Graph	Not named	P2	Yes	Privacy evaluation backbone (indicator/logic/case + data entity graph) and rule/indicator retrieval for scenario-based assessment; supports privacy-vs-availability trade-off analysis and gating	Privacy protection effectiveness scoring; operational decision support (trade-off + export gating + anonymization prompt)	N.A.	N.A.
Stade et al. (2025)	2025	E	ClimChg	PhysClimRisk	Other: Unspecified data source (see data_sources_summary)	Text; Tabular	Created: lacks ontology	P3	Unk	KG+ML	NER; RE; Summ; Retr; QA; Other	Ent:90230; Rel:70895	Sch:no; KG:partial; Data:partial; EP:3; https://environmentalevidence.org/ceeder-search/
KG-PLUB	2025	E	NatCap	BiodivLU	Standards: UN Sustainable Development Goals (SDGs); Web-scraped documents / web pages; Scientific literature / scholarly articles	Text	Not named	P3	Yes (Val,Part)	Semantic knowledge infrastructure for biodiversity-oriented pattern language; supports structured exploration and KG-grounded generation of design guidance	Biodiversity design recommendations via chatbot; KG query/exploration for education and design decision support; workshop-validated learning improvements	Ent:368; Rel:16	https://doi.org/10.57760/sciencedb.27653
Vanapalli et al. (2025)	2025	G	CorpCondInt	RegComp	Web-crawled regulatory websites/portals/legislative DBs + RSS feeds + news outlets; extracted outputs structured as JSON/CSV into KG	Text; Tabular	Not named	P1	Unk	Real-time, explainable policy backbone; metadata enrichment for LLM analytics; querying substrate for impact analysis	Monitoring + alerts, change summarization, deduplication, impact analysis, obligation/risk identification, and change prediction	N.A.	https://github.com/Kishorevb/policyinsight
Cai et al. (2025)	2025	H	ESGIntAss	ESGRate	ESG reports (PDF), multi-agency ESG ratings, derived sentence/action/SHAP data	Text; Tabular	Not named	P1	Yes(Onto)	Semantic and interpretability backbone for KG-augmented LLM ESG evaluation	ESG scoring, rating prediction, explainable analysis, and recommendation support	5 entity types; 5 relation types	N.A.
Chung et al. (2025)	2025	G	CorpCondInt	RegComp	GDPR regulatory text; unstructured scenario narratives; first-party legal decisions for labels	Text	Not named	P3	Yes	Structural reasoning scaffold for regulatory compliance, handling cross-references, exceptions, and scope before LLM judgment	Automated GDPR compliance checking and article-level violation detection	N.A.	N.A.
MPKG	2025	S	ProdRespCustSafe	ProdQual	Enterprise machining process texts and multi-modal industrial records (text/image/video) over 2013–2024; structured/semi/unstructured	Text; Tabular; Sensor / time-series; Image / raster; Event logs	Not named	P2	Yes (Val)	Graph-based fault knowledge base and reasoning scaffold embedded into LLM to improve inference and reduce hallucination	CNC machine fault diagnosis/localization and troubleshooting recommendation generation	N.A.	N.A.
QChsG	2025	S	ProdRespCustSafe	ProdQual	Quality standards, technical specifications, inspection reports (structured/semi/unstructured); RPV case data	Text; Tabular	Quality Characteristics Ontology Model (QChsOM) and Quality Formation Ontology Model (QFOM)	P2	Yes	Interpretable modeling of quality formation and importance analysis of QChs	KQChs identification and quality management decision support	Case-level: Ent:226; Trip:535	N.A.
Katzenstein and Etcheverry (2025)	2025	E	PolWasCiru	HazWaste	Sensor streams and third-party agencies; Uruguay MA data (2021–2023)	Tabular; Sensor / time-series; Geospatial	AIRQorg, AIRQreg, AIRQmed	P1	Yes	Provenance-aware, quality-aware air quality data management across the lifecycle		N.A.	https://gitlab.fing.edu.uy/air-data-quality/vocabularies-and-ontologies

doi:10.20944/preprints202602.1970.v2

Table A12. *Cont.*

Resource	Year	P	T	F	Data Source	Modality	Ontology	Cstr	Expert	KG Usage	Tasks	KG Scale	Availability
Jomraj et al. (2025)	2025	G	CorpCondInt	RegComp	Regulatory documents segmented into atomic sections (domain: regulated health/life sciences mentioned)	Text	Not named, A Schema	P4	Yes	Triplet-first retrieval + provenance grounding + subgraph visualization for traceable compliance QA	Regulatory compliance QA with auditable evidence	N.A.	N.A.
Remote Sensing Early Warning Knowledge Graph (RSEW-KG)	2025	E	ClimChg	PhysClimRisk	EM-DAT; sector websites; encyclopedias/news; GF-2 & Sentinel-1 imagery (attributes)	Text; Tabular; Image / raster; Geospatial	Remote Sensing Early Warning ontology/schema	P2	Yes	RS-driven, rule-based early warning and decision support	Automated warning issuance; disaster knowledge querying/visualization	N.A.	https://public.emdat.be/data
KG-IRDM	2025	S	HumCap	WorkManRel	Employee profiles, job descriptions, resumes, surveys, historical transfer records	Text; Tabular	Not explicitly named	P2	Yes (Annot,Val)	Semantic HR modeling and dynamic decision support	Job recommendation, personnel–position matching, HR supply–demand forecasting	N.A.	N.A.
Mishra et al. (2025)	2025	H	SDFrame	SDGs	Public unstructured data (news/social media/Wikipedia/blogs) + private curated datasets for NER/RE	Text	Not named	P2	Yes	Knowledge base to reason about SDG/climate indicator relationships and cascading effects of actions/policies	Climate-change decision support via KG-based relationship/cascade exploration; ontology/KG generation from text	N.A.	N.A.
TAIR (Trustworthy AI Requirements) knowledge graph	2025	G	CorpCondInt	AIEthics	EU AI Act text; international standards (ISO/IEC SC42, AI MSS templates)	Text	TAIR ontology	P1	Yes	Mapping standards conformance to regulatory compliance with traceability	Compliance mapping, gap analysis, regulatory decision support	N.A.	https://tair.adaptcentre.ie/; https://tair.adaptcentre.ie/demo.html
SDG-KG	2025	H	SDFrame	SDGs	Heterogeneous open data (RDB, NoSQL, APIs); UN SDG metadata; population datasets; OpenStreetMap	Tabular; Geospatial	SDG Graph; Metadata Graph	P3	Yes (Ont,Val)	Metadata-driven integration and automated SDG indicator computation with provenance	SDG indicator computation, conflict resolution, provenance-aware visualization	N.A.	https://github.com/wissalbenjira/sdg-kg
Wu et al. (2025)	2025	H	ESGIntAss	ESGRate	IFC BIM models (Revit/IFC/ifcopenshell), GIS web services, real-time sensor data (WAQI)	Text; Tabular; Sensor / time-series; Geospatial; CAD/BIM	Both: foundational ontology; Foundational ontology	P1	Yes (Part)	Semantic integration and ontology-driven reasoning to compute ESG indicators from GIS/BIM/IoT data	Real-time multi-scale ESG evaluation and interactive 3D urban planning visualization	N.A.	N.A.
LCI knowledge graph	2025	E	ClimChg	ProdFoot	KBOB LCI database; Bauteilkatalog component catalog (structured/semi-structured)	Tabular	Not named	P1	Yes	Semantic enrichment of LCI data and cross-level reasoning for embodied environmental impact	Early-stage building LCA and sustainable design decision support	N.A.	N.A.
Zhang et al. (2025)	2025	S	HumCap	OHS	198 iron & steel accident reports (2010–2024), public official/regulatory/media reports	Text	Not named	P3	Yes (Val)	RAG grounding + multi-hop causal reasoning scaffold to improve trustworthiness/interpretability	Safety QA, hazard identification, accident cause/root cause analysis, and safety recommendations	Ent:1637; Rel:4; Trip:2285	N.A.
SOCKG	2025	E	NatCap	BiodivLU	AgCROS experimental dataset (USDA); NALT for terminology alignment	Tabular; Geospatial	Not Named	P1	Yes (Ont,Val)	Semantic integration and large-scale querying of soil carbon experimentation data	SOC modeling support, treatment comparison, carbon sequestration analysis, decision support		https://idir.uta.edu/sockg/
Jin et al. (2025)	2025	E	NatCap	RawSrc	Industry research reports; online web content retrieved via search engines	Text	Not Mentioned	P3	Yes	Supply chain transparency enhancement and network-level analysis	Supply chain mapping, transparency assessment, and risk-related network analysis	300k nodes; 640k relationships	N.A.
Zheng and Brintrup (2025)	2025	S	HumCap	LabStanSC	Third-party supply chain dataset (automotive/MarkLines)	Tabular	Not named	P1	Yes	Factual anchor to ground GenAI embeddings and reduce hallucination in relationship prediction	Supply chain visibility via (contextual) relationship prediction (quintuplet prediction)	N.A.	N.A.
Zhu et al. (2025)	2025	G	CorpGov	FinRepQ	CSMAR financial database; supplier–customer disclosures; regulatory fraud labels	Tabular	Not Named	P1	Yes	Integrating financial data with supply chain structure for fraud detection and interpretation	Multi-year financial fraud detection and supply chain risk propagation analysis	Up to 44,611 nodes; >190k edges (10-year graph)	N.A.

doi:10.20944/preprints202602.1970.v2

Table A12. *Cont.*

Resource	Year	P	T	F	Data Source	Modality	Ontology	Cstr	Expert	KG Usage	Tasks	KG Scale	Availability
Tax Law Knowledge Graph (TaxKG)	2025	G	CorpCondInt	TaxTrans	IRC, Treasury Regulations, IRS guidance, court cases; expert-annotated UTP scenarios	Text	Not Named	P2	Yes (Val)	Structured legal reasoning, citation expansion, and explanation grounding	UTP risk classification and legally grounded explanation generation	Ent:50000; Trip:180000	N.A.
Tang and Chi (2025)	2025	G	CorpCondInt	FairComp	Zhihu antitrust discussions (2010–2021), 8,169 discussions	Text; Tabular	Not Named	P1	Unk	Network-structure + sentiment-propagation analysis of serialized corporate emergencies	Crisis/public opinion monitoring, risk node identification, regression-based factor analysis	N.A.	N.A.
Karjou et al. (2025)	2025	E	ClimChg	EnergyMix	IFC BIM data; IoT sensor streams; equipment data sheets	Text; Tabular; Sensor / time-series; Image / raster; CAD/BIM	Brick Schema; RealEstateCore Ontology	P1	Yes	Semantic integration and context-aware querying for smart-district energy performance	Energy assessment, semantic analytics, natural language building data access	Energy assessment, semantic analytics, natural language building data access	N.A.
Jain et al. (2025)	2025	S	IncSolSocAcc	DigIncl	Hexa-X-II enabler metadata, KPIs/KVIs, design principles, use-case specs	Tabular; Sensor / time-series	Not Named	P1	Yes	Explainable, sustainability-aware 6G E2E system design	Enabler selection, KPI/KVI alignment, sustainable system architecture design	N.A.	N.A.
Chen et al. (2025)	2025	E	NatCap	ResUseMatEff	Professional literature + statistical yearbooks (China/Changsha 2021)	Text; Tabular	Not Named	P1	Yes	Structured representation of iron/steel constituent mapping and mineralization lifecycle drivers for urban mining	Urban minerals constituent estimation support; lifecycle mineralization process modeling; sustainability/circular economy analysis	N.A.	N.A.
Uddin et al. (2025)	2025	S	IncSolSocAcc	HealAcc	Kaggle US Health Insurance dataset; Kaggle synthetic healthcare dataset (KG triples from medical-condition/treatment/test fields)	Tabular	Not Named	P2	Yes(Onto)	Semantic enrichment of patient representation for context-aware RL billing optimization	Cost-aware billing decision optimization with diagnostic accuracy retention	N.A.	https://www.kaggle.com/datasets/teertha/ushealthinsurancedataset; https://www.kaggle.com/datasets/prasad22/healthcare-dataset
W2RKG	2025	H	SysSustMod	CircLoop	4,499 Scopus papers (abstracts + review full texts)	Text; Tabular	A schema	P3	Yes (Annot, Prompt)	Scalable, standardized W2R knowledge base to support industrial symbiosis matching	IS opportunity identification (partner matching + network planning) and interactive exploration	3,518 waste entities; 4,471 resource entities; 33,679 W2R relations	https://github.com/nancycyzl/W2RKG-construction-with-LLMs; https://github.com/nancycyzl/W2RKG_application;
WHOW-KG	2025	E	NatCap	WaterSS	19 datasets (CSV/RDF) from ISPRA and ARIA/Lombardy plus controlled vocabularies	Tabular; Text; Sensor	WHOW ontology network (Hydrography, Water Monitoring, Water Indicator, Weather Monitoring, Health Monitoring; total 8 modules)	P1	Yes (Ont,Val,Part,Reg)	Distributed, FAIR semantic integration of water and health monitoring data	Cross-domain monitoring queries and decision support for water quality, health indicators, and extreme events	Trip:265453111	https://lod.dati.lombardia.it/sparql; https://doi.org/10.5281/zenodo.14510373
Hanžel et al. (2025)	2025	E	ClimChg	EnergyMix	20 open household electricity datasets + external socio-economic metadata	Tabular; Sensor / time-series; Geospatial; Graph	Not named	P1	Yes (Onto)	Semantic integration and large-scale analysis of residential electricity consumption	Energy analytics, forecasting, disaggregation, and policy support	Ent:791813; Rel:38; Trip:1577483	https://github.com/sensorlab/energy-knowledge-graph; https://sparqlelec.ijs.si/sparql
Wu et al. (2024)	2024	E	PolWasCiru	EWaste	Real end-of-life power battery pack; structural and disassembly knowledge	Tabular; Other	Not Named	P1	Yes(Onto)	Structured reasoning and optimization of disassembly sequences	End-of-life battery disassembly planning and recycling efficiency improvement	N.A.	N.A.
Tang et al. (2024)	2024	E	NatCap	WaterSS	Structured hydrological tables; scheduling rules and historical records; Pihe River Basin case data	Text; Tabular; Sensor / time-series	Not Named	P3	Yes	Automatic, flexible reservoir optimization modeling and dynamic scheduling decision support	Flood control optimization, multi-objective reservoir scheduling, operational instruction generation	N.A.	N.A.
Semantic Knowledge Graph of European Mountain Value Chains	2024	E	NatCap	BiodivLU	Expert textual documents; Excel summaries; Wikidata; OpenStreetMap; Eurostat GISCO	Text; Tabular; Image / raster; Geospatial; Graph	Narrative Ontology (NOnt)	P2	Yes (Val,Part)	Narrative-based semantic integration and cross-region knowledge discovery	Policy analysis, territorial comparison, semantic/geospatial querying, interactive storytelling	504k RDF triples; 454 value-chain subgraphs	https://doi.org/10.6084/m9.figshare.c.7098079
Osman et al. (2024)	2024	G	CorpGov	FinRepQ	BPMN process models, organizational models, ESG data/knowledge objects	Graph; Other	Not Named	P1	Yes (Part)	ESG knowledge management, semantic traceability, and enterprise-level reasoning	ESG-aware BPM analysis, traceability, and conceptual decision support	N.A.	N.A.

doi:10.20944/preprints202602.1970.v2

Table A12. *Cont.*

Resource	Year	P	T	F	Data Source	Modality	Ontology	Cstr	Expert	KG Usage	Tasks	KG Scale	Availability
Jing and Li (2024)	2024	S	ProdRespCust Safe	ProdQual	Multi-source agricultural supply chain data (inspection, production, processing, sales)	Text; Tabular	Not Named	P2	Yes	End-to-end agricultural product quality traceability and safety supervision	Traceability, quality problem localization, recall decision support	N.A.	N.A.
Kilanioti and Papadopoulos (2024)	2024	H	SDFrame	SDGs	UN Statistics Division SDG API (TypeScript/node.js scripts); harvested SDG dataset; SDG taxonomy/ontology as linked open data	Text; Tabular	SDG ontology	P2	Yes	Collab	Faster similarity search and distributed knowledge matching; claimed support for causal analysis/inference, influence measurement, explainable decisions/recommendations	N.A.	https://unstats.un.org/sdgs/indicators/indicators-list/
Fathi (2024)	2024	S	HumCap	LearnDev	Education-provider APIs; web data (ads, portals, government sites); external labor market KGs	Text; Tabular; Graph; Event logs	Not Named	P1	Yes	Semantic backbone, data enrichment, reasoning support, cold-start mitigation	Educational and career recommendation, reskilling/upskilling guidance, explainable decision support	N.A.	N.A.
Asset Life Cycle Knowledge Graph (ALC KG)	2024	E	ClimChg	ProdFoot	Enterprise CSV datasets (asset register, BMS, suppliers, application server); 4,000-asset case study	Tabular	Not Named	P1	Yes (Part,Val)	Semantic backbone and analytical substrate for asset life-cycle management and sustainability reporting	Asset querying, cost analysis, life-cycle cost and sustainability decision support	N.A.	N.A.
AttacKG+	2024	S	ProdRespCust Safe	DataSec	500 unstructured CTI reports; MITRE ATT&CK TTP matrix; STIX/D3FEND references	Text; Graph	Not named	P3	Yes (Annot,Val)	Semantic and temporal representation of cyber attacks, attack reconstruction, threat analysis	CTI parsing, technique identification, attack reconstruction, security decision support	Ent:20350; Rel:10175	https://anonymous.4open.science/r/CTKEG_Appendix-19DC/
Wang et al. (2024)	2024	E	NatCap	WaterSS	97,056 inspection records (text + images) from South-to-North Water Diversion Project	Text; Image / raster	Created: multimodal risk knowledge graph ontology	P2	Yes (Annot)	Multimodal risk integration, propagation analysis, and engineering safety assessment	Risk identification, diffusion analysis, and decision support in water diversion projects	Ent:550471	N.A.
Causal Quality-related Knowledge Graph (CQKG)	2024	S	ProdRespCust Safe	ProdQual	Aerospace manufacturing documents (defect surveys, inspections, maintenance reports)	Text; Tabular	Both: knowledge graph schema	P2	Yes (Val)	Causal reasoning, LLM augmentation, interpretability in quality analysis.	Root cause analysis and decision support for aerospace product manufacturing	N.A.	N.A.
CarbonKG	2024	E	ClimChg	GHG	ERP systems, process event logs, equipment/energy/material/personnel records	Tabular; Sensor / time-series; Event logs	Not Named	P1	Yes	Carbon traceability, integration, and predictive analysis	Carbon accounting, flow analysis, and emission prediction in complex manufacturing	N.A.	N.A.
Li et al. (2024)	2024	E	SustSolTech	GreenBldg	Construction standards documents; bridge construction scheme texts	Text	Created: utilized ontology	P1	Yes	Constraint and semantic backbone to guide LLM reasoning	Automated compliance checking of construction schemes	N.A.	N.A.
Marconnet et al. (2024)	2024	H	SysSustMod	CircLoop	PLM/MPM data (eBOM, assembly/disassembly plans), material prices, labor rates, expert DfX/DfCD knowledge	Tabular; CAD/BIM; Other	Not named	P1	Yes	Context-aware recommendation and decision support for sustainable design	Proactive circular disassembly design, sustainability assessment, and cost estimation	N.A.	N.A.
Boukhelifa and Merabet (2024)	2024	G	CorpCondInt	RegComp	Banking regulations (multi-jurisdictional texts); annotated regulatory corpus; organizational compliance data	Text; Tabular; Event logs	Not Named	P2	Yes (Annot,Val)	Compliance reasoning, dependency modeling, and regulatory risk analysis	Automated compliance checking, violation detection, and legal risk management in banking	N.A.	N.A.
Hofmeister et al. (2024)	2024	E	ClimChg	PhysClimRisk	Geospatial building data, property sales, population grids, weather sensors, flood monitoring APIs, infrastructure datasets	Tabular; Sensor / time-series; Geospatial; Graph	Existing: GeoSPARQL	P4	Yes	Cross-domain integration, dynamic risk reasoning, and evidence-based decision support	Flood impact assessment, real-time monitoring, scenario planning, infrastructure resilience analysis	N.A.	https://github.com/cambridge-cares/TheWorldAvatar
Sun et al. (2024)	2024	E	NatCap	WaterSS	28 irrigation districts basic info; reports/manuals from 7 management offices; inspection log short texts	Text; Tabular	Not named	P2	Yes	Cross-source integration, visualization, and decision-support for irrigation issue management	Inspection issue understanding, measure retrieval, and intelligent decision option generation	Ent:4255; Rel:11; Trip:14839	N.A.
Jiang et al. (2024)	2024	S	HumCap	OHS	Crane safety standards, 86 accident reports, real-time sensor data from a digital twin	Text; Sensor / time-series; CAD/BIM	Not Named	P1	Yes	Semantic reasoning and early-warning within a digital twin	Unsafe hoisting detection, safety alerting, and construction safety decision support	N.A.	N.A.
Al Akasheh et al. (2024)	2024	S	HumCap	WorkManRel	IBM HR Analytics tabular dataset (1,470 employees)	Tabular	A schema	P2	Yes(Eval)	Relational feature learning to enhance ML prediction	Employee turnover prediction and explainable HR analytics	Ent:1470	N.A.

doi:10.20944/preprints202602.1970.v2

Table A12. *Cont.*

Resource	Year	P	T	F	Data Source	Modality	Ontology	Cstr	Expert	KG Usage	Tasks	KG Scale	Availability
Du et al. (2024)	2024	E	PolWasCiru	HazWaste	Government hazardous chemical incident reports; regulatory knowledge (conceptual)	Text	Not Named	P2	Yes (Ont)	Semantic integration, incident analysis, and safety decision support	Incident retrieval, profiling, statistical analysis, and hazardous chemical safety management	N.A.	N.A.
Gupta et al. (2024)	2024	H	SDFrame	ReptStd	GRI standards; FinSim4-ESG taxonomy; 4,331 Reuters sustainability news articles	Text	Both: GRI sustainability reporting framework; FinSim4-ESG 2022 shared task ESG ...	P2	Yes (Annot,Val)	ESG-aware retrieval and grounding for LLM-based QA	ESG news QA, investor insight generation, sustainability analysis	N.A.	N.A.
ESG-KG	2024	H	ESGIntAss	ESGRate	850k ESG-related Dow Jones news articles (1980–2022)	Text; Tabular	Not named	P2	Yes (Ont,Annot,Val)	Large-scale ESG discourse analysis and monitoring	Trend analysis, media monitoring, ESG research support	4M entities	N.A.
ESG Metric Knowledge Graph (ESGMKG)	2024	H	SDFrame	ReptStd	ESG reporting frameworks (IFRS, TCFD, TNFD, SASB); ESG datasets and organisational data	Text; Tabular	Both: proposed ontology; standard ontology	P1	Yes (Ont,Val,Part)	Semantic integration and decision support for ESG metric management	ESG reporting, metric alignment, compliance, and investment decision support	4M entities	N.A.
Wen (2024)	2024	S	HumCap	OHS	CSTR simulation data; operation logs; inspection records; 47 HAZOP sheets	Text; Sensor / time-series; Event logs	Not Named; risk ontology	P2	Yes(Onto)	Safety knowledge representation and proactive decision support	Risk identification, causal analysis, and proactive safety management	Ent:176; Trip:260	N.A.
Da Silveira et al. (2024)	2024	S	ProdRespCust Safe	ChemSafe	ECHA REACH, EPA CTD, NIOSH toxicology databases	Text; Tabular	Not Named	P1	Yes(Onto)	Integrated chemical knowledge access and LLM grounding	Chemical risk retrieval, healthcare decision support, natural-language querying	N.A.	N.A.
Extreme Climate Architecture Knowledge Graph	2024	E	ClimChg	PhysClimRisk	Polar architecture case databases; COMNAP data; OwnThink, CN-DBpedia; Antarctic environmental datasets	Text; Tabular; Graph	Created: DBpedia	P3	Yes	Knowledge organization, retrieval, and LLM grounding for architectural decision support	KGQA, design support, visualization, and early-stage planning for extreme climate architecture	Ent:432; Trip:1491	N.A.
Zhou et al. (2024)	2024	H	SDFrame	ReptStd	GRI and ESRS reporting standards; prior sustainability KG subgraphs	Text; Graph	RSO (Reporting Standards Ontology)	P3	Yes (Annot,Val)	Indicator alignment, retrieval enhancement, and semantic interoperability	Mapping indicators across sustainability reporting standards	Ent:52; Trip:214	https://github.com/OntoSustain/RSO
FSFD-TLKG	2024	G	CorpGov	FinRepQ	Structured financial statements (74 companies, 2009–2022) and regulatory fraud labels	Text; Tabular	Created: detection ontology	P1	Yes	Explainable fraud reasoning and pattern extraction	Financial statement fraud detection and regulatory decision support	N.A.	N.A.
Wang et al. (2024)	2024	E	SustSolTech	GreenBldg	IFC-based BIM models (ARC & MEP), geometric relations, equipment manuals, schedules	Text; Tabular; Sensor / time-series; CAD/BIM	Brick Schema	P1	Yes	HVAC topology reasoning and BIM-to-BEM integration	Automated building energy model generation and performance simulation	N.A.	N.A.
Human–Cyber–Physical Knowledge Graph	2024	S	ProdRespCust Safe	ProdQual	Expert knowledge, algorithms/models, sensor and production process data	Text; Tabular; Sensor / time-series; Image / raster; Other	Not Named	P2	Yes	Integrated reasoning and decision support in manufacturing quality control	Quality monitoring, diagnosis, and intelligent quality-control decision making	Case-level: >500 entities	N.A.
Vasiliu et al. (2024)	2024	E	ClimChg	ClimFin	Historical stock and commodity time-series; fundamental, technical, ESG, and sentiment indicators	Tabular	Not named	P1	Semantic integration and correlation-aware synthetic data generation	Extreme-scale synthetic financial time-series generation and algorithm testing	N.A.	https://graph-massivizer.eu/	
Raikar and Deepak (2024)	2024	H	ESGIntAss	ESGRept	Journalism datasets, Google Knowledge Graph, Web 3.0 metadata, eBooks, glossaries	Text; Tabular; Graph	Both: particular ontology	P2	Yes (Val)	Semantic enrichment and population for sustainable journalism	KG population, sustainability-focused journalism analytics, knowledge reuse	N.A.	N.A.
Bronzini et al. (2024)	2024	H	SDFrame	ReptStd	Sustainability reports (124 companies); ESG categorization; Refinitiv ESG scores	Text; Tabular	Existing: PROV	P3	Yes	ESG disclosure analytics and interpretability	ESG analysis, similarity studies, and ESG score explanation	Trip:40000	https://github.com/saturnMars/derivingStructuredInsightsFromSustainabil
Li et al. (2024)	2024	E	ClimChg	ProdFoot	Engineering lifecycle data; CPCD and other emission factor databases; construction standards	Tabular	Not Named	P1	Yes	Carbon footprint traceability and inventory completeness	Carbon footprint accounting and lifecycle-based carbon-reduction support	N.A.	N.A.
SILVANUS Knowledge Graph	2024	E	ClimChg	PhysClimRisk	IoT sensors, climate/weather data, forestry data, social media sensing, standard fire ontologies	Text; Sensor / time-series; Geospatial; Graph	SILVANUS Wildfire Management Ontology	P1	Yes (Ont,Val,Part)	Semantic data fusion and decision support in wildfire management	Wildfire monitoring, crisis response, risk assessment, and operational decision support.	N.A.	https://silvanus-project.eu/results/resources/ontology/;https://silvanus-project.eu/
Human-Centered Knowledge Graph (HCKG)	2024	S	HumCap	JobQual	Manufacturing PPR data, operator data, IoT/sensor data, industrial standards	Tabular; Sensor / time-series; Event logs	Not Named	P1	Yes (Val)	Human-centered reasoning, analytics, and decision support in Industry 5.0	Collaboration assessment, KPI analysis, resource allocation, and operator support	N.A.	https://github.com/abonyilab/HCKG

doi:10.20944/preprints202602.1970.v2

Table A12. *Cont.*

Resource	Year	P	T	F	Data Source	Modality	Ontology	Cstr	Expert	KG Usage	Tasks	KG Scale	Availability
Quek et al. (2024)	2024	E	SustSolTech	GreenBldg	BIM, GIS, IoT sensors, simulations, spreadsheets	Text; Tabular	Modular TWA ontologies (OntoBuiltEnv, OntoBIM, OntoCityGML, etc.)	P4	Yes	Interoperability, dynamic reasoning, and knowledge discovery in built environments	Urban energy analysis, digital twins, laboratory automation, decision support	N.A.	https://github.com/cambridge-cares/TheWorldAvatar
Shi and Wu (2024)	2024	S	HumCap	OHS	Unstructured elevator accident report texts (102 used for extraction/KG); 50 labeled reports for training; additional accident cases used for statistical risk analysis are mentioned but not clearly as KG inputs	Text	Elevator Safety Accident Ontology Model	P2	Yes (Val)	Structuring accident knowledge + retrieval; deriving data-driven causal influence matrix to support DEMATEL/ISM/MICMAC risk analysis and decision support	Risk-factor analysis (DEMATEL), hierarchical modeling (ISM), driving/dependence classification (MICMAC), prevention/control decision support	Ent:1829; Trip:2918	N.A.
Peng et al. (2024)	2024	E	ClimChg	ProdFoot	Ecoinvent 3.7 LCA database (structured); enterprise LCI tables (semi-structured); translation APIs for thesaurus	Tabular	Both: established ontology	P1	Yes	Semantic representation, background data recommendation, and LCA automation	Flow/process recommendation, automated LCI/LCIA, decision support for environmental impact assessment	Ent:22968; Rel:3; Trip:41479	N.A.
Li et al.	2024	S	HumCap	OHS	3,500 unstructured OSHA construction accident reports	Text; Tabular	Not named	P2	Yes (Val)	Grounding and validating LLM-based safety risk identification	Construction safety risk identification, hazard recognition, decision support	N.A.	N.A.
Usmanova and Usbeck (2024)	2024	H	ESGIntAss	DualMat	Unstructured corporate sustainability reports (14 EU-based companies)	Text	OntoSustain (extended)	P3	Yes (Annot,Val)	Structured ESRS representation, reasoning, and gap identification in sustainability reporting	ESRS compliance structuring, disclosure analysis, gap detection, ESG transparency	N.A.	N.A.
Benjira et al. (2024)	2024	H	SDFrame	SDGs	Web open data (French census; sport facilities) + UN SDG metadata	Text; Tabular	SDG schema from UN metadata	P3	Yes	Semantic integration and computation of SDG indicators	SDG indicator calculation, schema mapping, sustainability analytics	N.A.	N.A.
Hu et al. (2024)	2024	S	ProdRespCust Safe	DataSec	Unstructured OSCTI reports (security vendors/news/blogs) + MITRE ATT&CK	Text	Both: knowledge graph schema	P3	Yes	Integrating IoCs, entities, and TTPs for interpretable threat analysis	Threat hunting, attack attribution, intrusion analysis	Ent:50745; Trip:64948	https://github.com/Netsec-SJTU/LLM-TIKG-dataset
Mitra et al. (2024)	2024	S	IncSolSocAcc	FinAcc	Structured CMIE databases (ProwessIQ, Industry Outlook), Indian MSMEs (2016–2021)	Tabular	Created: knowledge graph schema	P1	Yes	Relational feature learning and enhanced credit risk assessment	MSME credit risk prediction, default classification	N.A.	N.A.
Machine Knowledge Graph (MKG)	2024	E	ClimChg	GHG	Enterprise BOMs, component metadata, qualified substitute component pairs	Tabular	Not Named	P2	Yes	Learning component similarity under non-homophily to improve Scope 3 emissions estimation	Substitute part identification, Scope 3 emissions calculation support	11,270 entities; 50,251 connectedTo and 1,613 similarTo relations	N.A.
Wen et al. (2024)	2024	E	NatCap	BiodivLU	National forest inventory reports (7th–9th), forestry yearbook/statistics, official bureaus (NBS/meteorology/forestry), China Weather Network air-quality data	Tabular	Not named	P1	Yes	Visual correlation analysis and interpretation of forest–environment relationships; policy insight	Visualization dashboards, correlation analysis, trend forecasting, decision support	N.A.	N.A.
Wu et al. (2024)	2024	E	ClimChg	PhysClimRisk	Literature (CNKI), social media (Weibo, Zhihu), disaster statistics, Sentinel-1/2 remote sensing, meteorological/seismic data	Text; Sensor / time-series; Image / raster; Geospatial	Not named	P3	Yes	Semantic integration, spatio-temporal reasoning, and automated monitoring orchestration	Landslide monitoring, deformation analysis, impact assessment, disaster decision support	Ent:106	N.A.
Zhou et al. (2024)	2024	S	HumCap	OHS	Safety regulations, construction safety reports, on-site construction images	Text; Image / raster	Not named	P3	Yes	Multimodal safety knowledge integration, reasoning, and querying	Safety incident analysis, causal reasoning, decision support for power grid construction safety	N.A.	N.A.
MAKG	2024	S	HumCap	OHS	581 China MSA accident reports; auxiliary maritime corpora for model pretraining	Text	Not named	P2	Yes	Knowledge storage, querying, reasoning, and decision support for maritime safety	Accident analysis, causal reasoning, pattern recognition, aggregation analytics, safety management	Ent:16090; Trip:20809	N.A.
Lu et al. (2024)	2024	H	SysSustMod	CircLoop	BIM (IFC/OmniClass), building standards, USEPA and local policies, environmental datasets	Text	DiCon ontology	P1	Yes (Val)	Semantic integration and reasoning for material recycling/reuse assessment	CD&W material evaluation, recycling/reuse decision support, circular economy planning	N.A.	N.A.
Ren et al. (2024)	2024	S	ProdRespCust Safe	ProdQual	Multimodal cold chain data—sensor readings, logistics records, text logs, and images	Text; Tabular; Image	Not named	P2	Yes	Semantic integration and real-time monitoring of cold chain product quality	Product quality traceability, anomaly detection, visualization, and decision support in cold chain logistics	N.A.	N.A.

doi:10.20944/preprints202602.1970.v2

Table A12. *Cont.*

Resource	Year	P	T	F	Data Source	Modality	Ontology	Cstr	Expert	KG Usage	Tasks	KG Scale	Availability
NRKG	2024	S	IncSolSocAcc	NutriHeal	Food.com recipe and review dataset; derived nutritional attributes	Tabular	Not named	P1	Yes	Nutrition-aware representation learning and recommendation	Personalized food recommendation; healthy diet promotion	N.A.	https://www.kaggle.com/code/aayushmishra1512/food-recommender
NW1	2024	E	ClimChg	GHG	Web/social: Wikipedia	Tabular; Geospatial; Graph	OntoEnergySystem, OntoPowSys, OntoEIP, OntoCAPE	P4	Yes	Interoperable, provenance-aware power system modelling and scenario reasoning	Power system simulation, decarbonisation analysis, SMR deployment planning, energy policy decision support	N.A.	https://github.com/cambridge-cares/TheWorldAvatar
OfficeGraph	2024	E	ClimChg	EnergyMix	Real-world IoT sensor logs (444 devices, 17 models), building metadata, Wikidata	Sensor	SAREF	P1	Yes (Ont,Part)	Interoperable IoT data integration, analytics, and ML	Building management analytics, sustainability monitoring, graph-based ML experiments	Trip:89599577	https://github.com/RoderickvanderWeerdt/OfficeGraph; https://zenodo.org/records/10245815; https://data.interconnect.labs.vu.nl
Gilliard et al. (2024)	2024	S	ProdRespCustSafe	DataSec	CVE/CWE/CAPEC/CPE, Snort IDS alerts, AIT log dataset, system configuration scans	Graph; Event logs	Both: for ontology	P1	Yes	Semantic integration and automated reasoning for cyber attack detection	Intrusion detection, attack pattern inference, cybersecurity decision support	N.A.	N.A.
PGD-KG	2024	E	SustSolTech	GreenBldg	Parametric design models, regulatory codes, sustainable design literature, expert semantic templates	Text; CAD/BIM; Other	PGD-KG schema	P1	Yes	Knowledge-informed reasoning to prune solution space and accelerate PGD optimization	Sustainable building generative design, compliance checking, performance evaluation, multi-objective optimization	N.A.	https://github.com/GeorgeZWu/PGD_KG_Schema
Power Equipment Management Using Knowledge Graph	2024	H	ESGIntAss	ESGRept	Standards: TCFD Recommendations; Corporate sustainability / ESG reports; Internal enterprise data (logs, maintenance, operations)	Text; Tabular	Not Named	P1	Yes	Semantic integration and management of ESG metrics for reporting	ESG metrics extraction, standardized sustainability reporting, ESG data management	N.A.	N.A.
PrivComp-KG	2024	S	ProdRespCustSafe	DataSec	GDPR text (chunked to vector DB), vendor privacy policies (OPP-115), regulatory/obligation knowledge encoded as KG instances	Text; Graph	Not named	P3	Yes	Compliance inference and gap analysis between privacy policies and regulatory requirements	GDPR compliance verification, missing-article detection, policy improvement support	N.A.	https://github.com/<anonauthor>/PrivComp-KG.git
RSOKG	2024	H	SDFrame	ReptStd	GRI and ESRS standards; one real-world sustainability report; reused ontologies (ORG, QUDT, SKOS, DCMI)	Text	Sustainability Reporting Standards Ontology (RSO)	P1	Yes (Ont,Part)	Semantic interoperability and indicator mapping between sustainability reporting standards	N.A.		https://github.com/OntoSustain/RSO
Greif et al. (2024)	2024	E	ClimChg	ProdFoot	Ökobau.dat LCI database, expert domain knowledge, LLM-generated synthetic data, 3DP case data	Text; Tabular	Not named	P1	Yes	Integrated LCA reasoning and early-stage sustainability decision support	LCA analysis, emission calculation, scenario comparison, sustainable design optimization	Ent:135; Rel:128	N.A.
Supply Chain Knowledge Graph (SC-KG)	2024	S	ComRigRisks	RiskSrc	Structured supply chain databases (Marklines, Achilles)	Tabular; Graph	Not named	P1	Yes	Neurosymbolic reasoning and hidden risk discovery	Supply chain risk management, hidden dependency inference, complex risk querying	N.A.	N.A.
Kurniawan et al. (2024)	2024	S	ProdRespCustSafe	DataSec	Local logs/events/infrastructure info + public threat intel (CVE/CWE/CAPEC/MITRE ATT&CK, etc.); log use case uses AIT-derived dataset	Text; Tabular; Graph; Event logs	Created: Unified Cybersecurity Ontology (UCO); UCO	P3	Yes	Knowledge-grounded RAG context retrieval (graph queries + embedding-based similarity) to support cybersecurity analysis and reduce ungrounded responses	Threat-intel QA, CVE/vulnerability assessment over updated data, and security log analysis QA	N.A.	https://w3id.org/sepses/
KGSCS	2024	S	ProdRespCustSafe	HealVuln	Questionnaires, smart device sensor data, nursing publications, open-source medical knowledge	Text; Tabular; Sensor / time-series	Not named	P1	Yes (Annot,Val,Part)	Semantic integration, reasoning, risk identification, and personalized care support	Care report generation, KG-based QA, risk analysis, caregiver decision support	N.A.	N.A.
Pang et al. (2024)	2024	E	SustSolTech	ReEneg	Technical documents, project reports, databases, logs, and real-time PV sensor data	Text; Tabular; Sensor / time-series; Event logs	Not named	P2	Yes (Val)	Knowledge management, data fusion, querying, and report automation	Automated engineering reports, PV project analytics, visualization, decision support	N.A.	N.A.
Androna et al. (2024)	2024	E	ClimChg	PhysClimRisk	Eurostat, World Bank, OECD, WHO, Copernicus, EEA, UN datasets, policy documents	Text; Tabular; Sensor / time-series; Image / raster; Geospatial	Existing: SDGs	P1	Yes	Interoperable climate vulnerability knowledge management and analysis	CCVA, indicator aggregation, risk mapping, sensitivity analysis, policy support	N.A.	https://gitlab.com/netmode/sustaingraph

doi:10.20944/preprints202602.1970.v2

Table A12. *Cont.*

Resource	Year	P	T	F	Data Source	Modality	Ontology	Cstr	Expert	KG Usage	Tasks	KG Scale	Availability
UrbanKG	2024	H	SDFrame	SDGs	UN SDG database, EU SDG indicators (Eurostat), Copernicus, policy documents	Text; Tabular; Sensor / time-series; Geospatial	Existing: SDGs	P1	Yes	Semantic integration and analysis of SDG interlinkages	Synergy/trade-off analysis, SDG network analysis, multi-scale sustainability assessment	N.A.	https://gitlab.com/netmode/sustaingraph
VulKG	2024	S	ProdRespCustSafe	DataSec	NVD, CVE Details, CWE, Exploit Database (EDB)	Text; Tabular	Not named	P1	Yes	Structural vulnerability reasoning and risk assessment	Co-exploitation discovery, vulnerability prioritization, link prediction	Ent:276676; Trip:833456	https://github.com/happyResearcher/VulKG.git
Xiao and Sun (2023)	2023	S	IncSolSocAcc	DigIncl	Shanghai government open data platforms, elderly care service platforms, user profiles	Text; Tabular; Other	Not named	P1	Yes (Val)	Semantic storage and reasoning over elderly care policies and services	Personalized policy recommendation, care information QA, senior-friendly decision support	N.A.	N.A.
Emission Conversion Factors Knowledge Graph (ECF KG / CFKG)	2023	E	ClimChg	GHG	OpenKB/ont: Wikidata; ECF	Tabular; Graph	ECFO — Emission Conversion Factor Ontology; PECO — Provenance of Emission Calculations Ontology	P1	Yes (Ont,Val)	Transparent, machine-understandable emissions accounting and provenance	Carbon footprint calculation, provenance explanation, validation, ML emissions reporting	Ent:42400; Trip:662992	https://w3id.org/tec-toolkit; https://github.com/TEC-Toolkit/cfkg; https://github.com/TEC-Toolkit/Data-Validation
Wang et al. (2023)	2023	E	NatCap	WaterSS	Academic papers; EV manufacturers' and battery suppliers' reports	Text; Tabular	Not named	P2	Yes	Structural reasoning and decision support for robotic disassembly	Disassembly sequence planning, recycling optimization, operator guidance	N.A.	N.A.
Wu et al. (2023)	2023	E	ClimChg	GHG	NOAA climate data, AviationStack flight data, SimpleMaps cities, Climatiq emissions API	Tabular; Sensor / time-series; Geospatial	Not named	P1	Yes	Semantic integration and analytics of climate–tourism interactions	Tourism analytics, climate-aware recommendations, emissions analysis, travel decision support	N.A.	https://github.com/futaoo/climate-tourism-kg
Ameri et al. (2023)	2023	E	NatCap	RawSrc	Simulated CSV and JSON datasets based on grain elevator–processor scenarios; IFT CTE/KDE framework	Tabular; Event logs	Supply Chain Traceability Ontology (SCT)	P1	Yes (Ont)	Semantic traceability, reasoning over custody/ownership, and interoperability	Food traceability, contamination investigation, compliance, and supply-chain analytics	N.A.	N.A.
Oladeji and Mousavi (2023)	2023	E	ClimChg	GHG	Earnings call transcripts; Bill of Lading shipping records; synthetic emissions values	Text; Tabular	Both: knowledge graph schema	P3	Yes (Annot,Unk)	Conceptual modeling and analysis of supply-chain E-liability flows	Carbon accounting, auditing, benchmarking, policy analysis	N.A.	N.A.
Wang et al. (2023b)	2023	E	NatCap	WaterSS	Emergency plan documents (unstructured	semi-structured) and structured monitoring data	Not named	P2	Yes	Knowledge integration and intelligent emergency decision support	Emergency plan recommendation, water diversion risk management	N.A.	http://www.mwr.gov.cn/
Ma et al. (2023)	2023	E	ClimChg	GHG	Bridge construction documents, organization designs, carbon emission analysis reports	Text; Tabular	Not Named	P2	Yes (Val)	Knowledge integration and low-carbon decision support in bridge construction	Construction scheme recommendation, carbon-aware comparison, green construction planning	N.A.	N.A.
Peng et al. (2023)	2023	E	ClimChg	ClimFin	Tianjin carbon market platforms (unstructured; paper also notes possible structured DB + semi-structured logs/JSON + unstructured HTML/reports)	Text	Not Named	P2	Yes	Crawled text (news announcements/other info) from Guangdong & Data integration/management + Neo4j-based visualization, semantic querying, and cross-region comparison	Comparative analysis and visualization; semantic query	Trip:30047	N.A.
Stagnol et al. (2023)	2023	E	SustSolTech	CleanTech	IEA reports; GDELT 3.0 global news; custom topical document collections (unstructured/semi-structured)	Text; Tabular; Graph	Both: knowledge graph schema	P2	Yes (Annot)	Knowledge reconstruction, semantic linkage, novelty and sentiment monitoring	QA, novelty detection, market intelligence, decision support	N.A.	N.A.
Company Knowledge Graph	2023	G	CorpGov	OwnCtrl	Regulatory / enterprise ownership databases (structured)	Tabular; Graph	Not named	P1	Yes	Reactive reasoning, control detection, simulation	Company control analysis, regulatory supervision, what-if simulations	Ent:8589000; Trip:7749000	N.A.
Construction Accident Knowledge Graph (CAKG)	2023	S	HumCap	OHS	OSHA accident reports (2017–2021); real-time site BBS observations	Text; Tabular; Event logs	Not named	P1	Yes (Ont,Val,Part)	Objective risk quantification, dynamic safety analysis	Construction risk assessment, key behavior identification, safety decision support	1,543 accident records; 104 entities; 458 relations; 30 BBS indicators	N.A.
Wang and El-Gohary (2023)	2023	S	HumCap	OHS	OSHA 29 CFR 1926 fall-protection regulations	Text	Not named	P2	Yes (Annot)	Semantic representation and reasoning over safety requirements	Automated compliance checking, violation detection, safety analytics	Trip:7927	N.A.
Li et al. (2023)	2023	S	ProdRespCustSafe	DataSec	NYT10, SemEval-2010, WikiData5M	Text	Not named	P2	Unk	Semantic representation and management of sensitive personal information	Privacy protection, sensitive data analysis, KG construction automation	N.A.	N.A.
DaanMatch Knowledge Graph (DaanKG)	2023	H	SDFrame	SDGs	UN SDG metadata; NGO relational databases; scraped NGO web data; Wikidata geography; media evidence	Text; Tabular; Geospatial; Graph; Other	DaanKG Ontology (SDG, NGO, Geography sub-ontologies)	P1	Yes	Semantic integration, reasoning, CSR compliance and impact analysis	CSR–NGO matching, auditing, monitoring, decision support	N.A.	https://github.com/mdebellis/Daan_Knowledge_Graph

doi:10.20944/preprints202602.1970.v2

Table A12. *Cont.*

Resource	Year	P	T	F	Data Source	Modality	Ontology	Cstr	Expert	KG Usage	Tasks	KG Scale	Availability
Aravind Krishnan and Deepak (2023)	2023	S	HumCap	DEI	DEI document datasets; NELL; DBpedia; Google KG API; web-crawled data	Text; Graph	Not named (DEI ontology)	P2	Yes	Semantic integration, reasoning, inclusive DEI knowledge modeling	Ontology generation, DEI analytics, policy and research support	N.A.	N.A.
E-Liability Knowledge Graph	2023	E	ClimChg	GHG	Proposed use of corporate reports, logs, sensors, news, social media	Text; Tabular; Sensor / time-series; Image / raster; Event logs	Not named	P2	Yes	Carbon accounting transparency and E-liability tracking	Supply-chain emissions accounting, liability management, policy support	N.A.	N.A.
Ji et al. (2023)	2023	E	ClimChg	ReEneg	Dispatch cloud systems, D5000 alarms, fault logs, regulations, historical records	Text; Tabular; Sensor / time-series; Event logs	Created: fault-handling ontology	P2	Yes	Semantic reasoning and real-time fault-handling decision support	Fault diagnosis, rule-based reasoning, dispatch decision assistance	Ent:160000; Rel:547000	N.A.
Food4healthKG	2023	S	IncSolSocAcc	NutriHeal	FoodData Central, FoodOn, Chinese Food Ontology, KEGG, NCBI, MENDA, MiKG, MeSH, SNOMED CT	Text; Tabular; Graph	Both: food-centered ontology; Biomedical ontology	P1	Yes (Val)	Semantic integration, reasoning, nutrition intelligence	Food recommendation, mental health decision support	Ent:1637915; Rel:140; Trip:13346991	https://github.com/ccszbd/Food4healthKG
I-KNOW-FOO knowledge graph	2023	E	ClimChg	PhysClimRisk	ECOCROP, FoodOn, FIO, NEVO, FAOSTAT, synthetic pricing data	Tabular; Graph	Not named	P1	Yes (Ont)	Semantic reasoning on climate-resilient, low-CO_2 diets	Food substitution, sustainable diet decision support	N.A.	https://git.wur.nl/FoodInformatics/i-know-foo.git
Kebede et al. (2023)	2023	H	SysSustMod	CircLoop	Conceptual product lifecycle and compliance data	Other	Not named	P1	Yes	Conceptual support for Digital Product Passports and circular economy	CE decision support, lifecycle transparency, DPP implementation	N.A.	N.A.
GENA	2023	S	IncSolSocAcc	NutriHeal	PubMed abstracts; DOID, CHEBI, FOODON, MFOMD, APADISORDERS, ASDTTO, PR, FMA, SYMP; BC5CDR	Text; Graph	Not named	P2	Yes (Annot,Val)	Semantic integration and discovery of nutrition–mental health knowledge	Knowledge retrieval, relation discovery, biomedical research support	Ent:28598; Trip:43367	https://github.com/ddlinh/gena-db
Wang et al. (2023)	2023	E	PolWasCiru	EWaste	EV battery papers, reports, manufacturer/supplier documents	Text; Tabular	OWL-based EV battery ontology	P2	Yes	Semantic modeling and reasoning for robotic disassembly	Disassembly sequence planning, recycling optimization	N.A.	N.A.
Papageorgiou et al. (2023)	2023	E	NatCap	WaterSS	Industrial process knowledge, sensor data, enterprise systems, simulation outputs	Tabular; Sensor / time-series; Other	Not named	P1	Yes	Semantic integration and querying of water treatment process knowledge	Knowledge retrieval, KPI analysis, process optimization support	N.A.	N.A.
IoT-Reg Knowledge Graph	2023	S	ProdRespCustSafe	DataSec	GDPR, HIPAA, NISTIR 8228; reused IoT and privacy ontologies	Text; Sensor / time-series; Geospatial; Graph	IoT-Reg ontology	P1	Yes	Semantic integration and reasoning for IoT privacy compliance	Real-time compliance checking, risk mitigation, privacy decision support	N.A.	N.A.
Aprilia et al. (2023)	2023	E	PolWasCiru	PackWaste	Consumer opinion text (questionnaire responses)	Text	Both: knowledge graph schema	P2	Yes (Curate)	Semantic exploration of consumer cognition in sustainable packaging	Concept discovery, brainstorming, design decision support	N.A.	N.A.
Job Hazard Analysis Knowledge Graph (JHAKG)	2023	S	HumCap	OHS	115 JHA documents; 12 Codes of Practice; expert interviews	Text	O-JHAKG	P1	Yes (Ont,Val,Part)	Semantic reasoning and automation of JHA knowledge	Hazard identification, risk evaluation, control-measure planning	N.A.	N.A.
KG4NH	2023	S	IncSolSocAcc	NutriHeal	PubMed literature; FoodData Central; KEGG; MENDA; DMDA; FoodOn; SNOMED-CT	Text; Tabular; Graph	Not Named	P2	Yes (Ont,Annot,Val)	Semantic integration and reasoning for nutrition–health knowledge	KGQA, diet recommendation, clinical and research support	Ent:7437819; Rel:154; Trip:255017496	https://github.com/ccszbd/KG4NH
Regulatory Knowledge Graph	2023	G	CorpCondInt	RegComp	ADGM regulatory documents (COBS, rulebooks)	Text	Existing: PROV	P2	Yes (Ont,Annot,Val,Reg)	Explainable, executable compliance automation	Compliance checking, rule interpretation, decision support	231k nodes; 1.2M relations	https://github.com/Vladimir-Ershov/adgm-kg1
KG4NH	2023	S	IncSolSocAcc	NutriHeal	FDC, KEGG, NCBI Taxonomy, SNOMED-CT (structured ontologies) + biomedical literature (unstructured; text-mined associations)	Text; Tabular; Graph	Not named	P2	Yes	Knowledge reasoning / predicting reliable relations (triple existence) over nutrition–microbe–disease KG	Link prediction / missing relation completion (predicting new knowledge)	Ent:2367; Rel:3; Trip:65082	
Mandilara et al. (2023)	2023	H	SDFrame	SDGs	Policy documents (EGD, CSRs), OSDG dataset, SDG indicators, third-party data	Text; Tabular	Not Named	P2	Yes	SDG tracking, policy analysis, data enrichment	Policy–SDG mapping, analytics, decision support		https://gitlab.com/netmode/sdg-detector
Ospan et al. (2023)	2023	E	NatCap	WaterSS	Sensors, regulations, national statistics, PDFs, web tables, Wikipedia	Text; Tabular; Sensor / time-series; Geospatial	Not named	P1	Yes	Semantic integration, reasoning, decision support, spatiotemporal analysis	Water quality monitoring, pollution assessment, health correlation, sustainability planning	N.A.	https://github.com/Titrom025/PyTableMiner/tree/main/ontology
Peng et al. (2023)	2023	S	HumCap	OHS	Standards: Regulatory / policy documents; Incident / safety / accident reports	Text; Tabular	Created: predefined ontology; comprising ontology	P2	Yes	Semantic integration and decision support for hazard management	Hazard querying, analysis, and control-measure recommendation	328 hazard entities; 776 relations	N.A.
Salary Knowledge Graph	2023	S	HumCap	JobQual	University salary records (2013–2022)	Tabular	Not Named	P1	Yes	KBQA, analytics, reasoning, optimization	Salary QA, ranking, prediction, allocation decision support	N.A.	N.A.
LCIKG	2023	E	ClimChg	ProdFoot	Ecoinvent LCI datasets (UPR, cumulative LCI, product systems)	Tabular	Not named	P1	Yes	Semantic LCI data management, querying, interoperability	LCI retrieval, LCA analysis, activity comparison, sustainability decision support	Ent:40000000; Trip:100000000	N.A.

doi:10.20944/preprints202602.1970.v2

Table A12. *Cont.*

Resource	Year	P	T	F	Data Source	Modality	Ontology	Cstr	Expert	KG Usage	Tasks	KG Scale	Availability
MetaKG	2023	S	HumCap	LearnDev	Enterprise talent–course logs, skill profiles, organizational relations; Last.FM for evaluation	Tabular	A schema	P1	Yes	Context modeling and explainability in recommendation	Explainable course recommendation, CTR prediction, cold-start recommendation	Ent:30279; Rel:13; Trip:2039676	N.A.
Nature FIRST KG	2023	E	NatCap	BiodivLU	EUNIS, IUCN, Natura 2000, Corine Land Cover; RDF, CSV/XLS, shapefiles	Text; Tabular; Geospatial; Graph	Created: SiteOntology	P1	Yes (Annot)	Semantic integration, geospatial reasoning, FAIR biodiversity data backbone	Biodiversity preservation, wildlife movement prediction, recommender systems	Ent:371411	https://sensingclues.poolparty.biz/
NHANES Knowledge Graph	2023	S	IncSolSocAcc	HealAcc	NHANES survey data, codebooks, data dictionaries (2013–2018)	Tabular; Graph	Created: employing ontology; employed ontology	P1	Yes	Semantic data integration and equity-focused health analytics	Disparity analysis, equity measurement, policy-oriented healthcare access evaluation	N.A.	http://nhanes.eci.ufmg.br:9000/hadatac
Observatory Knowledge Graph (OKG)	2023	S	ComRigRisks	HRDD	Twitter posts on inequality; NLP-derived annotations (entities, dependencies, rolesets)	Text; Graph	OBservatory Integrated Ontology (OBIO)	P2	Yes	Semantic backbone for fine-grained, explainable discourse analysis	Social media observatories, inequality discourse analysis, policy-relevant insights	9.24M triples; 1.08M entities	https://github.com/muhai-project/okg_media_discourse https://api.druid.datalegend.net/datasets/lisestork/OKG/services/OKG/sparql
Liu and Zhan (2023)	2023	S	ProdRespCustSafe	DataSec	Unstructured CTI reports; rule-extracted IoCs	Text	Both: Unified Cybersecurity Ontology (UCO); CTI-specific ontology	P3	Yes	Not named	CTI extraction, attack knowledge modeling, cybersecurity analysis	N.A.	N.A.
Wang et al. (2023a)	2023	E	NatCap	WaterSS	Water diversion emergency plan documents (2014–2021)	Text	Not Named	P2	Yes	Emergency knowledge integration and decision support	Emergency entity/relation extraction, KG construction, emergency response support	Trip:790000	https://www.kaggle.com/datasets/lihuwang/ptm-mfgcn
Zhao (2023)	2023	E	ClimChg	GHG	Electric-vehicle patent texts	Text; Tabular	Not Named	P2	Yes	Technology value quantification and carbon credit calculation	Carbon emission technology valuation, carbon trading decision support	N.A.	N.A.
Yan et al. (2023)	2023	E	NatCap	WaterSS	Web pages (Wikipedia, Baidu), scientific literature (CNKI)	Text; Sensor / time-series	Both: water ontology	P2	Yes (Annot)	Knowledge-enhanced feature weighting and correlation modeling	Water quality prediction, parameter importance learning	Ent:468; Trip:340	N.A.
Simone et al. (2023)	2023	S	HumCap	OHS	Near-miss reports from Seveso industrial establishments	Text	Near-miss safety ontology	P2	Yes (Val)	Safety meta-analysis and completeness assessment	Near-miss reporting evaluation, safety management decision support	Ent:45000; Rel:75000	N.A.
ScrapKG	2023	H	SysSustMod	CircLoop	ISRI standards, UNS, proprietary company scrap data, textual scrap descriptions	Text; Tabular	Not Named	P1	Yes	Semantic integration and ML-driven inference for circular economy	Scrap classification, material identification, circular economy decision support	N.A.	N.A.
Liu et al. (2023)	2023	E	NatCap	RawSrc	Siemens internal supply data; public/private customs data; media data	Text; Tabular	Created: knowledge graph schema	P2	Yes	Transparency, reasoning, and resilience analysis	Link prediction, critical supplier detection, supply-chain risk management	Ent:65277; Rel:11; Trip:311676	N.A.
He et al. (2023)	2023	E	NatCap	WaterSS	Maps, national agencies, water bulletins, historical records, expert knowledge	Text; Tabular; Geospatial	Not named	P2	Yes	Semantic integration and smart recommendation	Water-use policy recommendation, probability-based decision support	Ent:200; Rel:9	N.A.
Wang et al. (2023)	2023	E	ClimChg	EnergyMix	Smart meter electricity data (450 households, 2021–2022)	Tabular; Sensor / time-series	Both: and ontology	P1	Yes	Safety-oriented knowledge integration and explainable analytics	Unsafe electricity prediction, early warning, household similarity analysis	N.A.	N.A.
Supplier–Customer Knowledge Graph	2023	G	CorpGov	FinRepQ	CSMAR supplier–customer disclosures and financial statements (2017–2020)	Tabular	Not named	P2	Yes	Relational enrichment and graph-based fraud reasoning	Financial statement fraud detection, supply-chain risk analysis	Ent:3921; Rel:2; Trip:6681	N.A.
Le Guillarme and Thuiller (2023)	2023	E	NatCap	BiodivLU	FunFun, BETSI, GloBI (CSV/API); taxonomic references	Tabular; Other	NCBITaxon; Soil Food Web Ontology (SFWO)	P1	Yes	Semantic integration and ecological reasoning	Multitrophic analysis, trophic group inference, soil food-web studies	N.A.	https://zenodo.org/record/1216257) ; https://github.com/nleguillarme/inteGraph
Food Safety Temporal Knowledge Graph	2023	S	ProdRespCustSafe	ProdQual	National food safety sampling data (China, 2018–2021)	Tabular	Created: food ontology; core ontology	P1	Yes	Temporal modeling and interpretable food risk prediction	Food risk forecasting, hazard prediction, regulatory decision support	Ent:19732; Rel:6; Trip:143021	N.A.
Energy Knowledge Graph	2023	E	SustSolTech	ReEneg	SCADA systems, MySQL time-series, weather and plant metadata	Tabular; Sensor / time-series; Geospatial	Not Named	P1	Yes	Semantic interoperability, explainable analytics, energy data spaces	Smart grid analytics, energy monitoring, service integration	18.3M triples	N.A.

doi:10.20944/preprints202602.1970.v2

Table A12. *Cont.*

Resource	Year	P	T	F	Data Source	Modality	Ontology	Cstr	Expert	KG Usage	Tasks	KG Scale	Availability
Global EEE Green Design Knowledge Graph	2023	E	PolWasCiru	EWaste	Global standards, regulations, certifications (PDF/HTML/XLS/CSV/XML)	Text; Tabular	Created: Reference ontology	P1	Yes	Objective, comprehensive green design evaluation	Index weighting, green degree calculation, product comparison, design decision support	Ent:6300; Trip:22000	N.A.
Enterprise Knowledge Graph	2023	G	CorpCondInt	TaxTrans	Enterprise credit data, tax arrears records, trust-breaking events, macroeconomic statistics	Text; Tabular	Not named	P1	Yes	Feature extraction and modeling of tax-risk contagion	Tax arrears prediction, enterprise risk management	Ent:2845112; Rel:142945345; Trip:142945345	N.A.
Yan et al. (2023)	2023	E	NatCap	WaterSS	Structural representation and feature learning for water usage analysis	Tabular; Sensor / time-series	A schema	P1	Yes (Ont)	Structural representation and feature learning for water usage analysis	Water usage clustering, anomaly detection, resource management decision support	N.A.	N.A.
Alam and Ali (2022)	2022	S	IncSolSocAcc	FinAcc	Home Credit Default Risk dataset	Tabular	The Financial Industry Business Ontology (FIBO)	P1	Yes	Feature enrichment and explainable credit-risk modeling	Loan default prediction, credit risk management	1.63M nodes; 12 relationship types	N.A.
Loh et al. (2022)	2022	S	HumCap	LearnDev	Skills Framework, job descriptions, surveys, SGA interview data	Text	Not Named	P1	Yes (Ont)	Guided reasoning and explainable conversational control	Skills gap analysis, workplace learning, talent development	N.A.	N.A.
AttacKG	2022	S	ProdRespCust Safe	DataSec	CTI reports; MITRE ATT&CK procedure examples	Text	MITRE ATT&CK	P2	Yes	Aggregating and enriching attack technique knowledge	Technique identification, attack reconstruction, variant detection, cybersecurity analysis	N.A.	https://github.com/li-zhenyuan/Knowledge-enhanced-Attack-Graph
Wu et al. (2022)	2022	G	CorpGov	FinRepQ	CSMAR audit data (2013–2019) + audit opinion reports; web-crawled historical names for disambiguation	Text; Tabular	Both: knowledge graph schema	P1	Yes	Path-search reasoning to find potential fraud corporations and mine interpretable fraud features	Fraud risk analysis / detection; audit-feature mining via reasoning paths	Ent:2980; Rel:7; Trip:6934	N.A.
Sharma et al. (2022)	2022	E	ClimChg	GHG	Web-crawled climate/carbon articles + Our World in Data CO2 datasets	Text; Tabular	Existing: SDGs	P2	Yes	Semantic integration, querying, reasoning, and enrichment for carbon footprint analysis	Carbon footprint analysis, semantic search, policy and decision support	N.A.	N.A.
Tauqeer et al. (2022)	2022	G	CorpCondInt	RegComp	Digitally modeled contracts (structured RDF); manually provided contract instances	Tabular; Graph	smashHitCore ontology	P1	Yes (Val)	Semantic representation and reasoning for GDPR contract compliance	Contract compliance verification, auditing, violation detection, decision support	N.A.	https://github.com/AmarTauqeer/Contract/tree/master/backend/
CSKG4APT	2022	S	ProdRespCust Safe	DataSec	OSCTI reports; CTI standards (STIX, CAPEC, CVE, NVD); expert-curated corpora	Text; Tabular	Not Named	P2	Yes	Threat knowledge integration, reasoning, and attribution	APT attribution, threat hunting, cybersecurity decision support	Ent:2608327; Rel:7; Trip:12935201	N.A.
Pedro et al. (2022)	2022	S	HumCap	OHS	OSHA accident dataset (structured CSV); 200 curated construction cases	Text; Tabular	Construction Safety Ontology	P1	Yes	Semantic sharing, integration, and analytics of construction safety information	Accident retrieval, safety knowledge sharing, trend analysis, decision support	Case-level KG with 200 accident instances	https://github.com/lanrepedro3/constructionsafetyontology
Janev et al. (2022)	2022	E	ClimChg	EnergyMix	SCADA/MySQL, IoT/PMU, meteorological JSON/XML, ENTSO-E; external KGs	Tabular; Sensor / time-series; Graph	Not named	P1	Yes	Interoperable, responsible knowledge management in energy data ecosystems	Forecasting, balancing, predictive maintenance, analytics, decision support	N.A.	N.A.
Weichselbraun et al. (2022)	2022	S	HumCap	LearnDev	Education provider websites (unstructured) + curated occupation knowledge base	Text; Tabular; Graph	Both: education ontology	P2	Yes (Ont,Val)	Semantic integration and enrichment for reskilling/upskilling intelligence	Career path recommendation, education recommendation, semantic search	Ent:73969; Trip:734447	https://github.com/fhgr/careercoach2022
Risk Knowledge Graph in Railway Safety (RKGRS)	2022	S	HumCap	OHS	British railway accident/incident reports (427 documents, GOV.UK)	Text; Tabular	Created: infrastructure ontology; topology ontology	P2	Yes	Causal modeling and quantitative risk assessment	Hazard identification, risk evaluation, safety decision support	N.A.	N.A.
ESG Knowledge Graph	2022	E	ClimChg	ClimFin	ESG reports, sustainability reports, public ESG datasets	Text; Tabular	Not Named	P2	Yes	Transparent ESG evaluation and green premium reduction	ESG scoring, green premium ranking, persuasive AI decision support	N.A.	N.A.
Diseasomics knowledge graph	2022	S	IncSolSocAcc	HealAcc	Biomedical ontologies; Wikipedia/PubMed/textbooks; EHR comorbidity data; DisGeNET; PharmGKB	Text; Tabular; Graph	Both: The Symptom Ontology (SYMP); Human Disease Ontology (DO)	P2	Yes (Annot,Val)	Machine-interpretable disease knowledge and clinical reasoning	Differential diagnosis, decision support, knowledge discovery	Ent:75642; Trip:6292931	https://triage.cyberneticcare.com/diseasePrediction; https://zenodo.org/doi/10.5281/zenodo.6416938
Liang et al. (2022)	2022	E	NatCap	WaterSS	Statistical yearbooks; hydrological/meteorological/socio-economic inputs	Tabular; Geospatial	Not Named	P1	Yes	Visual orchestration and decision support for water regulation	Dynamic, multi-scenario water resources regulation and simulation	N.A.	http://www.yunqishui.com/pages/g/gamePost.shtml?view=true&postId=298&u=99999
Knowledge Graph of Dangerous Goods (KGDG)	2022	E	PolWasCiru	EnvComp	Web-crawled DG descriptions; standard DG reference documents	Text	Existing: SDGs	P2	Yes	Semantic representation and inference of DG knowledge	DG safety management, transport/storage decisions, emergency support	N.A.	N.A.
SustainGraph	2022	H	SDFrame	SDGs	UN SDG, Eurostat, NDCs, policy documents, climate	Text; Tabular; Sensor /	SustainGraph Ontology	P2	Yes	Semantic integration, interoperability, reasoning,	SDG monitoring, policy alignment, nexus analysis,	N.A.	https://gitlab.com/netmode/

doi:10.20944/preprints202602.1970.v2

Table A12. *Cont.*

Resource	Year	P	T	F	Data Source	Modality	Ontology	Cstr	Expert	KG Usage	Tasks	KG Scale	Availability
Wang et al. (2022)	2022	S	HumCap	JobQual	CVs, job descriptions, Wikipedia, Google Maps	Text; Geospatial; Graph	Not Named	P2	Yes (Val)	Semantic reasoning and optimization for staffing recommendations	Job matching, cross-domain staffing recommendation, location-aware decision support	N.A.	N.A.
KnowUREnvironment	2022	E	ClimChg	PhysClimRisk	Scientific article abstracts from S2ORC	Text	Created: Standardized ontology; concept ontology	P2	Yes	Explainable climate knowledge representation and reasoning	QA, IR, recommendation, fact-checking, knowledge discovery	Ent:10321; Rel:4323; Trip:24263	https://github.com/saiful1105020/KnowUREnvironment
Wang et al. (2022)	2022	E	ClimChg	ProdFoot	LCA databases (Ecoinvent), enterprise systems, scientific literature, equipment manuals	Text; Tabular	Created: process-flow ontology; through ontology	P2	Yes (Ont,Annot,Val)	Knowledge-enriched LCA reasoning and inventory data support	LCI recommendation, process–flow identification, impact indicator selection	N.A.	N.A.
Manager Knowledge Graph	2022	G	CorpGov	FinRepQ	Manager employment data; CSMAR financial databases (China A-share, 2011–2017)	Tabular	Not Named	P1	Yes	Generating topological features capturing managerial interlocks	Financial fraud detection and classification using ML	Trip:473918	N.A.
OpenBiodiv biodiversity knowledge graph	2022	E	NatCap	BiodivLU	Biodiversity literature; RI databases (GBIF, CoL, ELIXIR, ENA); repositories (Zenodo, PMC)	Text; Tabular; Image / raster; Geospatial; Graph	OpenBiodiv-O; TaxPub; EML	P2	Yes (Val,Part)	Semantic integration, FAIR data linking, interoperability	Data discovery, integration, analytics support	N.A.	N.A.
LinkClimate knowledge graph	2022	H	ESGIntAss	ESGScen	NOAA climate APIs; OpenStreetMap; Wikidata	Tabular; Sensor / time-series; Geospatial; Graph	Climate Analysis (CA) ontology	P1	Yes (Val)	Semantic integration, data enrichment, interoperable climate analysis	Climate querying, time-series analysis, cross-domain data exploration, decision support	Trip:14000000	https://github.com/futaoo/LinkClimate
PRIVAFRAME	2022	S	ProdRespCustSafe	DataSec	DPV/DPV-PD taxonomy; FrameNet; WordNet; Framester	Text; Graph	PRIVAFRAME	P1	Yes	Explainable, logic-based representation of sensitive personal data	Sensitive information detection; fine-grained privacy analysis; hybrid symbolic–neural privacy protection	N.A.	https://w3id.org/framester/dpv2fn
Han et al. (2022)	2022	E	PolWasCiru	HazWaste	Survey reports, government documents, statistical yearbooks, web encyclopedias, vector maps, remote sensing data	Text; Tabular; Image / raster; Geospatial	Both: multi-facets ontology; generalized ontology	P1	Yes (Ont,Annot,Unk)	Semantic integration, reasoning, visualization, environmental risk analysis	Information retrieval, knowledge reasoning, pollutant pattern discovery, decision support	20,000 entities, 10,000 relations	https://github.com/Feng-David/Tranform-the-information-of-sites-into-neo4j.git
Yang et al. (2022)	2022	S	ProdRespCustSafe	ProdQual	Structured quality reliability databases; unstructured failure/quality documents; domain dictionaries	Text; Tabular	Created: reliability ontology; portable ontology	P2	Yes	Semantic integration, reasoning, and decision support in aviation quality reliability	Smart Q&A, entity retrieval, failure cause analysis, corrective measure recommendation, decision support	N.A.	N.A.
Diamantini et al. (2022)(Ongoing)	2022	H	SDFrame	ReptStd	ESG standards documents (DVFA-EFFAS, GRI)	Text	KPIOnto	P1	Yes	Semantic reference, reasoning, interoperability, and comparability of ESG metrics	Indicator calculation, dependency reasoning, consistency checking, ESG comparison and harmonization	N.A.	https://kdmg.dii.univpm.it/kpionto/specification/
Li et al. (2022)	2022	E	ClimChg	EnergyMix	Building configuration metadata; BAS operational sensor time-series data	Tabular; Sensor / time-series	Both: assessment ontology; detection ontology	P1	Yes	Semantic representation, reasoning, and fault diagnosis support	Fault detection, fault localization, interpretability and commissioning support	N.A.	N.A.
SEARCH-KG	2022	S	ProdRespCustSafe	ChemSafe	WISER, PubChem, CAMEO Chemicals, ICSCs, NIOSH; SMILES → MACCS fingerprints	Text; Tabular	Not named	P1	Yes	Semantic integration, reasoning, verification, real-time analytics support	Chemical diagnosis from symptoms, emergency response decision support	Ent:3263; Rel:171; Trip:196631	N.A.
Delgoshaei et al. (2022)	2022	E	SustSolTech	GreenBldg	Building automation data, equipment metadata, spatial, temporal, and unit information	Sensor	Multiple reused ontologies (Brick, SOSA, BOT, SAREF4BLDG, QUDT, Time) + core ontology	P1	Yes	Semantic integration, inference, and decision support	FDD, context-aware control, spatial & temporal reasoning, asset management, compliance	N.A.	N.A.
TERA	2022	S	ProdRespCustSafe	ChemSafe	ECOTOX, NCBI Taxonomy, EOL, PubChem, ChEMBL, MeSH, Wikidata	Tabular; Graph	Not named	P1	Yes	Semantic integration, background knowledge, embedding-based reasoning	Chemical adverse effect prediction, ecological risk assessment support	reduced KGC 241k triples, KGS 59k triples	https://doi.org/10.5281/zenodo.3559865; https://github.com/NIVA-Knowledge-Graph/TERA; https://github.com/NIVA-Knowledge-Graph/KGs_and_Effect_Prediction_2020
Koloski et al. (2022)	2022	H	SDFrame	TaxMap	Financial report sentences; WordNet; Wikidata5m	Text; Graph	Existing: PROV	P2	Yes	Knowledge-informed feature enrichment	ESG sustainability detection in short financial texts	N.A.	https://gitlab.com/boshko.koloski/formicca-finsem-esg
Water Knowledge Graph (WKG) and Water Information Network (WIN)	2022	E	NatCap	WaterSS	Sensors: Sensor / telemetry / spatiotemporal data	Tabular; Sensor / time-series	Not named	P1	Yes	Semantic modeling + embedding-based analytics to support grouping and decision-making for water management	Water zone quality classification; decision support/recommendation of corrective actions	N.A.	N.A.

doi:10.20944/preprints202602.1970.v2

Table A12. *Cont.*

Resource	Year	P	T	F	Data Source	Modality	Ontology	Cstr	Expert	KG Usage	Tasks	KG Scale	Availability
Gao and Liu (2021)	2021	G	CorpCondInt	EthicsAC	Standards: CARDS taxonomy	Tabular; Event logs	Not Named	P1	Yes (Ont,Part,Reg)	Semantic integration, anomaly analysis, and investigative analytics	Latent clue discovery, anomaly detection, community detection, core/key node identification, decision support	N.A.	N.A.
Wang et al. (2021)	2021	S	ProdRespCustSafe	ProdQual	Structured enterprise quality databases, audit and non-conformance records, domain and synonym thesauri	Text; Tabular	Not named	P1	Yes (Val)	Semantic integration, visualization, and decision support for aerospace product quality management	Querying, analytics, visualization, and quality-improvement decision support	Ent:12000; Rel:11	N.A.
Wang et al. (2021)	2021	E	ClimChg	EnergyMix	Remote signaling, telemetry measurements, component databases, simulated IEEE systems, real provincial grid data	Tabular; Sensor / time-series	Not Named	P1	Yes	Semantic representation and inference substrate for topology identification	Topology identification, fault-tolerant connectivity inference, island detection, operational decision support	N.A.	N.A.
Liu et al. (2021)	2021	G	CorpCondInt	EthicsAC	Unstructured disciplinary inspection notifications and reports from official supervision bodies	Text	Not Named	P2	Yes	Semantic organization, reasoning, and decision support in discipline inspection	Querying, relationship discovery, risk prevention, and anti-corruption support	Ent:27689; Trip:14909	N.A.
Zhang et al. (2021)	2021	E	PolWasCiru	EnvComp	Public government environmental websites, policy documents, enterprise and pollutant data	Text; Tabular	Created: knowledge graph schema	P2	Yes	Semantic integration and relational analysis to support enterprise forecasting	Enterprise development prediction, risk assessment, and decision support	Ent:20000	N.A.
Zhang et al. (2021)	2021	S	ProdRespCustSafe	ProdQual	Unstructured/semi-structured supervision texts, public opinion, inspection results, MSRA corpus	Text; Tabular	Not named	P2	Yes	Information integration and traceability across the supervision chain	Quality/safety supervision, product traceability, inspection analysis, regulatory decision support	Ent:66184; Trip:223640	N.A.
Zheng et al. (2021)	2021	S	ProdRespCustSafe	ChemSafe	National chemical catalogs; enterprise datasets; unstructured accident/procedure texts	Text; Tabular	Not mentioned	P2	Yes	Lifecycle-wide semantic integration and risk-management support	Knowledge retrieval, risk analysis, visualization, and QA preparation	Ent:66184; Trip:223640	N.A.
Wang et al. (2021)	2021	E	PolWasCiru	EnvComp	Environmental policy texts, public enterprise records, national industry classification standards	Text; Tabular	Existing: PROV	P2	Yes	Quantified semantic representation of policy–enterprise impacts	Enterprise-level policy impact prediction, risk analysis, and investment decision support	N.A.	N.A.
Financial Knowledge Graph (FKG)	2021	G	CorpGov	FinRepQ	PDF annual financial reports, semi-structured company data, structured financial databases	Text	Financial Ontology	P2	Yes	Semantic integration, querying, and financial storytelling	Financial QA, comparative analysis, decision support, report summarization	N.A.	N.A.
FoodKG	2021	S	IncSolSocAcc	NutriHeal	FoodKG (recipes, ingredients, nutrition); ADA health guidelines; Reddit-derived query patterns	Text; Tabular; Graph	Not named	P1	Yes	Constraint-aware reasoning and personalized retrieval	Personalized food recommendation, health-aware QA, dietary decision support	Trip:67000000	https://github.com/hugochan/PFoodReq
FSFD	2021	G	CorpGov	FinRepQ	RESSET financial statements; computed Pearson correlations	Tabular	Not Named	P2	Yes	Correlation-aware feature embedding	Financial statement fraud detection	N.A.	N.A.
Skills & Occupation Knowledge Graph	2021	S	HumCap	LearnDev	ISCO taxonomy, ESCO skills, 600k job postings, Textkernel Extract	Text; Graph	Not Named	P2	No (None)	Modeling labor-market skills demand and relationships	Skill-based matching, link prediction, career pathfinding, skill relevance analysis	Ent:1220; Trip:3910	N.A.
Mishra and Mittal (2021)	2021	E	ClimChg	PhysClimRisk	SciDCC dataset (11,539 Science Daily climate-change news articles)	Text	Both: knowledge graph schema	P2	Yes	Relationship analysis and reasoning on climate change factors	Climate-change KG construction, insight extraction, policy support	N.A.	https://sites.google.com/view/scidccdataset
Open-CyKG	2021	S	ProdRespCustSafe	DataSec	Unstructured APT and cybersecurity reports; MalwareDB and CTI datasets	Text	Created: cybersecurity ontology	P2	Yes	Cyber threat intelligence querying and analysis	Threat retrieval, relationship discovery, and cybersecurity decision support.	N.A.	https://github.com/IS5882/Open-CyKG
Xiaoping et al. (2021)	2021	E	ClimChg	EnergyMix	Dispatching-domain norms/standards/regulations + dynamic dispatching knowledge (document-based)	Text	Existing: RDF	P1	Yes	Knowledge organization + faster query/reasoning; supports knowledge learning/training/aux decision-making	Knowledge learning; skill training; online auxiliary decision-making	Ent:530; Rel:656	N.A.
Liu et al. (2021)	2021	S	ProdRespCustSafe	ProdQual	Thousands of quality problem records; dictionary data; unstructured descriptive texts	Text; Tabular	Created: knowledge graph schema	P1	Yes (Annot,Val)	Semantic knowledge backbone and retrieval/reasoning for quality management	Knowledge QA; decision support	N.A.	N.A.
Orellana and Piedra (2021)	2021	H	SDFrame	SDGs	CKAN open datasets; official UN SDG data; external linked open data sources	Text; Tabular; Graph	Not Named	P2	Yes	Semantic enrichment, interoperability, and knowledge integration for SDG open data	Semantic enrichment; linked data exploration; decision support	N.A.	N.A.

doi:10.20944/preprints202602.1970.v2

Table A12. *Cont.*

Resource	Year	P	T	F	Data Source	Modality	Ontology	Cstr	Expert	KG Usage	Tasks	KG Scale	Availability
Yoo et al. (2021)	2021	S	HumCap	ChemSafe	NIH WISER; PubChem; ChemSpider	Tabular; Other	Not named	P1	Yes	Reasoning and decision support in chemical hazard response	Chemical identification; emergency response guidance; symptom prediction for new chemicals	N.A.	N.A.
Aryan et al. (2021)	2021	E	ClimChg	EnergyMix	BIFROST simulation data (topology, measurements, actuations); external context (weather, customer data); expert causality knowledge	Tabular; Sensor / time-series; Other	ExpCPS knowledge graph	P1	Yes (Ont)	Explainability and causality reasoning in cyber-physical energy systems	Event explanation; root-cause analysis; decision support	N.A.	https://pebbie.org/expcps/
Gordon et al. (2021)	2021	S	ComRigRisks	ComRel	Manually curated REPD stakeholder data; reports, datasets, tools; external geospatial web services	Tabular; Geospatial	PPOp	P1	Yes (Ont,Annot,Part)	Semantic integration and coordination across multi-stakeholder environmental planning networks	Stakeholder discovery; coordination analysis; environmental decision support	N.A.	https://github.com/SDS-OKN/PPOp/
Janev et al. (2021)	2021	E	ClimChg	EnergyMix	SCADA and operational energy data; market platforms; meteorological data; external KGs (Wikidata/DBpedia)	Tabular; Sensor / time-series; Graph	Both: REFerence ontology; target ontology	P1	Yes	Semantic interoperability and controlled data exchange in energy data spaces	Energy balancing; forecasting; predictive maintenance; federated analytics	N.A.	N.A.
Eguiguren and Piedra (2020)	2020	H	SDFrame	SDGs	CKAN open data portals; UN SDG API; external LOD (DBpedia)	Text; Tabular; Graph	Not named	P2	Yes	Semantic integration and discoverability of SDG-related open data	Semantic integration and discoverability of SDG-related open data	Trip:2900000	https://jamesjose7.github.io/sdg-od/
Atzeni et al. (2020)	2020	G	CorpGov	OwnCtrl	Italian Chambers of Commerce enterprise ownership database (2005–2018)	Tabular; Graph	Existing: OWL	P1	Unk	Logic-based reasoning and augmentation of enterprise ownership knowledge	Company control; close-link/asset eligibility; family-link detection; financial supervision support	4.1M nodes, 4.0M edges per year	N.A.
Biospytial Knowledge Graph	2020	E	NatCap	BiodivLU	GBIF occurrences and taxonomy; WorldClim climate data; ETOPO1 DEM; IUCN Red List	Tabular; Image / raster; Geospatial; Graph	Existing: MeSH	P1	Yes	Semantic integration and spatial–taxonomic reasoning in ecology	Co-occurrence analysis; species distribution studies; biodiversity and conservation analytics	N.A.	https://github.com/molgor/biospytial
Fang et al. (2020)	2020	S	HumCap	OHS	Construction site images; safety codes; historical hazard reports	Text; Tabular; Image / raster; Event logs	Not Named	P1	Yes	Semantic reasoning and regulation-aware hazard detection	Automated hazard identification; FFH detection; safety compliance checking	N.A.	N.A.
Zhao et al. (2020)	2020	E	ClimChg	EnergyMix	Power grid operational data; textual power-network data processed via NLP	Text; Tabular; Sensor / time-series; Image / raster	Not named	P2	Yes	Visual analytics and semantic organization of power network data	Topology visualization; data mining; decision-support visualization	N.A.	N.A.
Beydon et al. (2020)	2020	S	IncSolSocAcc	FinAcc	Financial data lake; lending domain knowledge; external business ontologies	Tabular	Not Named	P1	Yes	Semantic data mediation, transfer learning, explainability in credit risk	Credit risk scoring; cross-domain transfer learning; synthetic data generation	N.A.	N.A.
Cybersecurity Knowledge Graph (CSKG)	2020	S	ProdRespCustSafe	DataSec	Vulnerability databases (CVE/CWE/NVD/CISA); security texts; ICS asset/network data	Text; Tabular; Graph	Created: and ontology	P2	Yes	Semantic integration and analysis of ICS cybersecurity threats	Threat correlation; vulnerability analysis; security decision support	3,878 relations (case-level EPIC KG); 19,838 training instances; 11 relation types.	N.A.
Power Knowledge Graph	2020	E	ClimChg	EnergyMix	Provincial electricity consumption; national industrial electricity statistics	Text; Tabular	Not Named	P2	Yes	Integration and querying of heterogeneous power statistics	Electricity consumption retrieval; visualization; data integration	N.A.	N.A.
Mao et al. (2020)	2020	S	HumCap	OHS	HAZOP reports; process diagrams; operational documents; relational databases	Text; Tabular; CAD/BIM	Created: constructed ontology; Methodology ontology	P2	Yes (Ont,Part)	Knowledge integration, reasoning, and decision support in process safety	Safety visualization; information retrieval; causal-chain extraction; emergency decision support	Ent:5615; Rel:2680	N.A.
Energy Knowledge Graph (EKG)	2020	E	SustSolTech	ReEneg	Energy standards (CIM, IEC 61850, EMIX), existing ontologies (IoT-ARM, OWL-S, SSN), scenario specifications	Text; Tabular; Sensor / time-series; Geospatial; Graph	Both: Suggested Upper Merged Ontology (SUMO); systematic ontology	P1	Yes (Ont,Part)	Semantic interoperability, reasoning, and integration of decentralized smart energy services	Microgrid modeling; energy trading; service composition; inference; complex event processing	N.A.	N.A.
Social-Impact Funding Knowledge Graph	2020	S	ProdRespCustSafe	RSInv	IRS 990 data; federal grants portals; official nonprofit/foundation sites; UN SDG documents	Text; Tabular	Both: both ontology	P2	Yes (Ont,Annot,Val,Part)	Semantic integration, reasoning, and public knowledge access for social-impact funding	Funding discovery; partner matching; SDG-aligned impact analysis; decision support	Ent:1046847; Rel:936419	N.A.
Lüdemann et al. (2020)	2020	G	CorpCondInt	TaxTrans	GLEIF; OECD tax rates; World Bank indicators; Wikidata	Tabular; Graph	Created: populated ontology	P1	Yes	Reasoning and anomaly detection in international taxation	Detect aggressive tax planning; identify anomalous ownership and address patterns	Ent:1587960; Trip:22839123	http://taxgraph.informatik.uni-mannheim.de/
XAI4Wind Knowledge Graph	2020	E	SustSolTech	ReEneg	Skillwind maintenance manual; SCADA data; alarm logs; fault labels; images	Text; Sensor / time-series; Image / raster; Event logs	Not Named	P1	Yes	Explainable, multimodal decision support in wind turbine O&M	Anomaly explanation; maintenance recommendation; interactive O&M decision support	Ent:537; Rel:9; Trip:1059	http://github.com/joyjitchatterjee/XAI4Wind

doi:10.20944/preprints202602.1970.v2

Table A12. *Cont.*

Resource	Year	P	T	F	Data Source	Modality	Ontology	Cstr	Expert	KG Usage	Tasks	KG Scale	Availability
Manufacturing Knowledge Graph (MKG)	2019	S	ProdRespCustSafe	ProdQual	Sensors: DBpedia; Wikipedia; Sensor / telemetry / spatiotemporal data; ...	Text; Tabular; Sensor / time-series; CAD/BIM	Both: structured ontology; STRUCTURED ontology	P1	Yes (Annot,Val,Part)	Semantic backbone for MK reuse, similarity-based ranking, navigation/matching, decision support in production problem solving	Problem query answering + solution recommendation; knowledge matching/navigation	Ent:4568; Rel:12658	http://58.206.100.146/fishbone/
Huang et al. (2019)	2019	S	IncSolSocAcc	HealAcc	Three Chinese healthcare websites (unstructured/semi-structured), Chinese Food Composition Table 2004, Chinese symptom KB	Text; Tabular; Graph	Not Named	P2	Yes (Ont)	Semantic integration, reasoning, and knowledge services for healthy diet management	Diet recommendation, semantic retrieval, intelligent QA, healthcare decision support	Ent:21443; Rel:5; Trip:281205	N.A.
Tang et al. (2019)	2019	E	ClimChg	EnergyMix	Multi-system enterprise data, operation & inspection records, equipment parameters, specifications, organizational info	Text; Tabular	Created: large-scale ontology; open-source ontology	P2	Yes	Semantic integration, traceability, and efficient equipment information retrieval	Equipment search, fault analysis, lifecycle management decision support	N.A.	N.A.
Eguiguren and Piedra (2019)	2019	H	SDFrame	SDGs	CKAN Open Data portals (CSV/XLS/XLSX), UN SDG textual descriptions	Text; Tabular	Existing: DCAT; RDF Data Cube; Dublin Core; SKOS; DBpedia	P2	Yes	Semantic integration and linking of Open Data to SDGs	SDG-oriented dataset discovery, visualization, and decision support	Ent:42924	N.A.
SEPSES Cybersecurity Knowledge Graph	2019	S	ProdRespCustSafe	DataSec	CVE, NVD, CVSS, CPE, CWE, CAPEC, Snort rules (CSV/XML/JSON)	Text; Tabular; Event logs	Not named (uses DCAT, Data Cube, SKOS)	P1	Yes	Semantic integration and linking of Open Data to SDGs	SDG-oriented dataset discovery, visualization, and decision support	Ent:644732; Trip:36594388	https://w3id.org/sepses;https://github.com/sepses/cyber-kg-converter;https://github.com/sepses/vocab
Sun and Wang (2019)	2019	S	ProdRespCustSafe	ProdQual	TE process simulation data (22 variables, 20 fault types) + prior knowledge	Tabular; Sensor / time-series	Created: knowledge graph schema	P1	Yes	Structural change detection and Bayesian fault reasoning	Multi-source fault detection and diagnosis; root-cause inference; efficiency comparison	N.A.	N.A.
Elluri et al. (2018)	2018	S	ProdRespCustSafe	DataSec	GDPR text, PCI DSS standard, CSA controls, cloud provider privacy policies	Text	Not Named	P2	Yes	Semantic representation and automation of data compliance	GDPR/PCI DSS compliance checking, policy validation, decision support	N.A.	N.A.
Jia et al. (2018)	2018	S	ProdRespCustSafe	DataSec	Vulnerability databases (CVE/NVD/CNVD/CNNVD etc.), security websites and forums, enterprise response centers, unstructured web text	Text; Tabular	Not named	P2	Yes	Semantic integration and deductive reasoning over cybersecurity knowledge	Entity extraction, knowledge deduction, intrusion detection support, situational awareness	N.A.	N.A.
Duan et al. (2018)	2018	S	ProdRespCustSafe	HealVuln	International eldercare guidelines; professional nursing literature	Text	Not Named	P1	Yes	Procedure representation and semantic integration of eldercare knowledge	Eldercare procedure generation, nursing decision support, caregiver guidance	N.A.	N.A.
Verona et al. (2018)	2018	S	HumCap	LabStanSC	Dirty List of Slavery, Brazilian electoral and donation data, NGO and religious organization datasets	Tabular	POLARE ontology	P1	Unk	Semantic integration and political power analysis	Power/influence measurement, SNA, investigative exploration and transparency analysis	1,853 nodes; 2,156 edges	http://www.trabalhoescravo.info
Energy Knowledge Graph (EKG)	2018	E	SustSolTech	ReEneg	Smart-grid standards (CIM, IEC 61850), IoT-ARM, OWL-S; domain knowledge for energy services	Other	Both: Generic ontology; Web ontology	P1	Yes	Semantic integration and interoperability of energy services	Micro-grid modeling, prosumer community formation, energy trading negotiation, service/process composition	N.A.	N.A.
Chi et al. (2018)	2018	S	IncSolSocAcc	NutriHeal	China Food Composition; Chinese health websites; TCM theory and culture	Text	Not Named	P2	Yes (Annot,Val)	Knowledge integration, semantic retrieval, diet education, recommendation	Healthy diet knowledge retrieval, personalized food recommendation, healthcare education support	N.A.	N.A.
Water Affair Knowledge Graph	2018	E	NatCap	WaterSS	Oracle monitoring tables; water affair texts; WordNet, DBpedia, CN-DBpedia, standards, expert knowledge	Text; Tabular; Graph	Created: Enterprise ontology; enterprise ontology	P2	Yes	Semantic integration and similarity-based reasoning	Water information recommendation, retrieval, and decision support.	N.A.	N.A.

 doi:10.20944/preprints202602.1970.v2

Appendix I. Glossary of Key Terms

This survey intersects ESG reporting practice, knowledge-graph engineering, and NLP/LLM system design. To keep terminology consistent across the **Data→KG** and **KG→App** stages, we group definitions into three tables: (i) ESG and reporting frameworks, (ii) KG and data-modeling concepts (including provenance/qualification), and (iii) NLP/LLM pipeline terminology (including agentic and validation-first workflows).

Table A13. Glossary of Key ESG, Reporting, and Sustainability Terms.

Term	Definition
ESG	**Environmental, Social, and Governance.** An umbrella framing for assessing sustainability- and ethics-relevant risks, impacts, and management practices, often operationalized through indicators and disclosures.
ESG-RFM	**ESG Research Focus Map.** The survey's vendor-agnostic **pillar–theme–focus** taxonomy (with crosswalks to major standards/ratings families) used to drive literature retrieval, coding, and analysis.
KG4ESG	The atlas-style survey and curated corpus of ESG/sustainability knowledge-graph works (2015–2025), organized as two coupled stages: **Data→KG** (construction) and **KG→App** (downstream use).
Materiality	A principle for determining which sustainability topics and indicators are relevant enough to report and manage. In practice, materiality criteria differ by audience (e.g., investors vs. broader stakeholders) and by reporting regime.
Double materiality	A reporting principle that considers both (i) **financial materiality** (how sustainability issues affect the firm) and (ii) **impact materiality** (how the firm affects people and the environment).
ESRS	**European Sustainability Reporting Standards.** A set of EU sustainability reporting standards that specify topical and cross-cutting disclosure requirements, including structured concepts such as materiality assessment and metric definitions.
CSRD	**Corporate Sustainability Reporting Directive.** EU legislation that expands and strengthens sustainability reporting requirements and is operationalized through ESRS-based disclosures.
GRI	**Global Reporting Initiative.** A widely used global standards family for sustainability reporting, emphasizing comparable disclosures on economic, environmental, and social impacts.
SASB	**Sustainability Accounting Standards Board (Standards).** Industry-oriented disclosure standards focusing on sustainability issues likely to be financially material, often used as an investor-facing materiality lens.
MSCI	**MSCI ESG Ratings (methodology).** A widely used ratings framework that structures ESG issues and evaluates company performance relative to peers, commonly used as a practical topic backbone in ESG analytics.
SDGs	**Sustainable Development Goals.** A set of 17 UN goals (2030 Agenda) used to frame cross-cutting sustainability targets and indicators across sectors and policy contexts.
TCFD	**Task Force on Climate-related Financial Disclosures.** A disclosure framework emphasizing climate-related risks/opportunities, governance, strategy, risk management, and metrics/targets (often tied to scenario analysis).
Scope 1–3 emissions	A categorization of GHG emissions commonly used in reporting: **Scope 1** (direct emissions from owned/controlled sources), **Scope 2** (indirect emissions from purchased energy), **Scope 3** (other indirect value-chain emissions).
LCA	**Life Cycle Assessment.** A methodology for quantifying environmental impacts across a product/service life cycle (e.g., cradle-to-gate or cradle-to-grave), requiring explicit assumptions, system boundaries, and units.
Greenwashing	Misleading claims or selective disclosures that create an unjustified impression of sustainability performance or alignment with standards (often motivating provenance- and evidence-linked verification).

Table A14. Glossary of Key Knowledge Graph, Representation, and Data-Management Terms.

Term	Definition
Knowledge Graph (KG)	A structured knowledge base representing entities as nodes and relationships as edges, typically governed by a schema/ontology; used for integration, querying, reasoning, and evidence-backed analytics.
Ontology	A formal specification of a domain conceptualization (classes, relations, constraints, and intended meaning). In KG4ESG, ontologies support standards-bound semantics and consistent indicator interpretation.
Schema	The structural blueprint that defines allowable entity types, relations, attributes, and constraints for a KG (including expected qualifiers such as time/unit/scope).
RDF	**Resource Description Framework.** A standard graph data model where facts are represented as triples (subject–predicate–object), enabling interoperable knowledge representation.

 doi:10.20944/preprints202602.1970.v2

Table A14. *Cont.*

Term	Definition
OWL	**Web Ontology Language.** A standard for expressing ontologies with formal semantics (e.g., class hierarchies, property constraints), commonly used to support reasoning over RDF data.
SPARQL	A standard query language for RDF graphs, supporting pattern matching and joins over triples (often used for compliance-style retrieval, KGQA backends, and auditable querying).
SHACL	**Shapes Constraint Language.** A standard for validating RDF graphs against schema-like constraints (e.g., required properties, value types, cardinalities), aligning with validation-first KG pipelines.
Triple (s, p, o)	The fundamental KG fact unit: **s**ubject (entity), **p**redicate (relation/property), **o**bject (entity or literal value).
Qualification / qualifiers	Additional metadata required to make facts computable and comparable (e.g., **time validity**, **unit**, **boundary/scope**, **method**, **uncertainty**). Central for ESG indicators and event facts.
Provenance	Metadata capturing the origin and lineage of a claim (source document/span/table row/sensor product, extraction method, version), enabling auditability, re-evaluation, and dispute/appeal workflows.
Alignment / crosswalk	A mapping between semantically related concepts across standards, taxonomies, or schemas (e.g., linking ESRS/GRI/SASB/MSCI labels), often required for cross-framework normalization.
Spatiotemporal modeling	Representing time and location explicitly (intervals, events, geospatial features, trajectories) to support hazard/risk analytics, monitoring, and multi-resolution environmental reasoning.
Unit & quantity modeling	Representing numeric values together with units and quantity kinds so that indicators remain computable (conversion, aggregation, comparability) rather than merely retrievable.
FAIR Principles	Guiding principles stating that data should be **F**indable, **A**ccessible, **I**nteroperable, and **R**eusable; often used to motivate open KG artifacts and standardized metadata.

Table A15. Glossary of Key NLP, LLM, and Graph-Enabled Pipeline Terms.

Term	Definition
Data→KG / KG→App	The survey's pipeline lens: **Data→KG** covers evidence-to-schema construction (typing, qualification, provenance); **KG→App** covers downstream use (retrieval, QA, verification, prediction, optimization) mediated by KG–NLP interfaces.
P1–P4 paradigms	Construction paradigm labels: **P1** ontology-first lifting and deterministic integration; **P2** rule/supervised extraction over a fixed schema; **P3** LLM-assisted structuring and alignment (with grounding/validation); **P4** tool-using/agentic pipelines with iterative validation and repair.
NLP	**Natural Language Processing.** Methods for analyzing and generating language, including extraction/linking, classification, retrieval, QA, summarization, and verification.
LLM	**Large Language Model.** A neural model trained on large text corpora for language understanding/generation; in KG4ESG often used for schema-conditioned extraction, alignment, KGQA, and evidence-backed synthesis.
Information Extraction (IE)	A family of tasks that converts unstructured text into structured representations, typically including **entity recognition**, **relation/event extraction**, and (often) **entity linking** to a KG schema.
Entity linking (EL)	Mapping text mentions to canonical entities/identifiers (and sometimes to schema types), enabling consistent aggregation across documents, disclosures, and registries.
RAG	**Retrieval-Augmented Generation.** A pipeline where a retriever selects evidence (documents/spans/KG subgraphs) that is then used as context for generation, improving factuality and traceability when evidence links are preserved.
GraphRAG	A RAG variant where retrieval is guided by graph structure (e.g., traversing multi-hop neighborhoods, relation-aware expansion), often used to assemble more coherent, entity-centered context for ESG queries.
KGQA / text-to-query	Question answering over KGs using structured queries (e.g., SPARQL) or graph-guided retrieval. Includes *text-to-SPARQL* and related interfaces that translate natural language into controlled KG access.
Agentic workflow	A multi-step controller (often LLM-driven) that plans actions, calls tools (retrievers, query engines, validators), and iteratively refines outputs; emphasized in P4 for traceability and reliability.
Tool use (tool calling)	Invoking external functions/systems (e.g., KG queries, calculators, rule checkers, geospatial operators) from within a pipeline; central for constraint enforcement and replayable audit traces.
Structured output	Producing machine-readable outputs under explicit formatting and type constraints (e.g., JSON schemas, ontology-aligned slots), used to reduce silent schema drift and ease validation.
Validation / constraint checking	Automatic checks that outputs conform to schema rules and ESG semantics (types, cardinalities, units, temporal scope, evidence links), often implemented with schema validators and domain rules.
Hallucination	A failure mode where a generative model produces plausible but unsupported statements or structure (e.g., invented triples, wrong units/scopes), motivating grounding and validation-first design.

 doi:10.20944/preprints202602.1970.v2

References

Friede, G.; Busch, T.; Bassen, A. ESG and financial performance: aggregated evidence from more than 2000 empirical studies. *Journal of sustainable finance & investment* **2015**, *5*, 210–233.

Li, H.; Abd Nikooie Pour, M.; Li, Y.; Lindecrantz, M.; Blomqvist, E.; Lambrix, P. A survey of general ontologies for the cross-industry domain of circular economy. In Proceedings of the Companion Proceedings of the ACM Web Conference 2023, 2023, pp. 731–741.

Pan, S.; Luo, L.; Wang, Y.; Chen, C.; Wang, J.; Wu, X. Unifying large language models and knowledge graphs: A roadmap. *IEEE Transactions on Knowledge and Data Engineering* **2024**, *36*, 3580–3599.

Gao, Y.; Xiong, Y.; Gao, X.; Jia, K.; Pan, J.; Bi, Y.; Dai, Y.; Sun, J.; Wang, H.; Wang, H. Retrieval-augmented generation for large language models: A survey. *arXiv preprint arXiv:2312.10997* **2023**, *2*.

Guo, T.; Chen, X.; Wang, Y.; Chang, R.; Pei, S.; Chawla, N.V.; Wiest, O.; Zhang, X. Large language model based multi-agents: A survey of progress and challenges. *arXiv preprint arXiv:2402.01680* **2024**.

Diamantini, C.; Khan, T.; Potena, D.; Storti, E.; et al. Shared Metrics of Sustainability: a Knowledge Graph Approach. In Proceedings of the SEBD, 2022, pp. 244–255.

Yu, M.; Rabhi, F.A.; Bandara, M. Ontology-driven architecture for managing environmental, social, and governance metrics. *Electronics* **2024**, *13*, 1719.

Markovic, M.; Garijo, D.; Germano, S.; Naja, I. TEC: Transparent emissions calculation toolkit. In Proceedings of the International Semantic Web Conference. Springer, 2023, pp. 76–93.

Yu, M.; Rabhi, F.; Xia, B.; Yang, Z.; Tan, F.; Lu, Q. OntoMetric: An Ontology-Guided Framework for Automated ESG Knowledge Graph Construction. *arXiv preprint arXiv:2512.01289* **2025**.

Cai, X.; Ma, M.; Hera, C.; Zulechner, B.; Fu, J.; Dawkins, J.; Lopez Uy, S.L.; Wang, X.; Wang, X. Evaluating Sustainability Rating Methodologies Using a Universal Dataset and AI Framework for Interpretive Integrity. *Available at SSRN 5267508* **2025**.

Ershov, V. A case study for compliance as code with graphs and language models: Public release of the regulatory knowledge graph. *arXiv preprint arXiv:2302.01842* **2023**.

Hofmeister, M.; Brownbridge, G.; Hillman, M.; Mosbach, S.; Akroyd, J.; Lee, K.F.; Kraft, M. Cross-domain flood risk assessment for smart cities using dynamic knowledge graphs. *Sustainable Cities and Society* **2024**, *101*, 105113.

Frakes, E.; Wu, Y.; French, R.H.; Li, M. GeoOutageKG: A Multimodal Geospatiotemporal Knowledge Graph for Multiresolution Power Outage Analysis. In Proceedings of the International Semantic Web Conference. Springer, 2025, pp. 221–239.

Diamantini, C.; Potena, D.; Rossetti, C.; Storti, E. A Knowledge Graph Framework for Impact Calculation in Life-Cycle Assessment **2025**.

Wu, T.; Li, J.; Bao, J.; Liu, Q.; Jin, Z.; Gao, J. CarbonKG: industrial carbon emission knowledge graph-based modeling and application for carbon traceability of complex manufacturing process. *Journal of Computing and Information Science in Engineering* **2024**, *24*, 081001.

Lippolis, A.S.; Lodi, G.; Nuzzolese, A.G. The Water Health Open Knowledge Graph. *Scientific Data* **2025**, *12*, 274.

Nananukul, N.; Kejriwal, M. HealthEQKG: A Knowledge Graph and Data Model for Health Equity Research. In Proceedings of the International Semantic Web Conference. Springer, 2025, pp. 183–202.

Wu, J.; Mayer, S.; Pilz, S.; Antille, Y.S.; Albert, J.L.; Stoll, M.; Garcia, K.; Fuchs, K.; Bally, L.; Eichelberger, L.; et al. FoodCoach: Fully Automated Diet Counseling. *IEEE Journal of Biomedical and Health Informatics* **2025**.

Bronzini, M.; Nicolini, C.; Lepri, B.; Passerini, A.; Staiano, J. Glitter or gold? Deriving structured insights from sustainability reports via large language models. *EPJ Data Science* **2024**, *13*, 41.

Garza, L.; Elluri, L.; Kotal, A.; Piplai, A.; Gupta, D.; Joshi, A. Privcomp-kg: Leveraging knowledge graph and large language models for privacy policy compliance verification. *arXiv preprint arXiv:2404.19744* **2024**.

Zhou, Y.; Gu, X.; Ding, J.; Chen, S.; Perzylo, A. Accessing the Capabilities of KGs and LLMs in Mapping Indicators within Sustainability Reporting Standards. In Proceedings of the Workshop on Natural Language Processing for Knowledge Graph Creation (NLP4KGC) at International Conference on Semantic Systems (SEMANTICS), 2024.

Boukhelifa, Y.; Merabet, R. Evaluating the Role of Natural Language Processing in Automating Regulatory Compliance and Legal Risk Management in the Banking Sector. *Studies in Knowledge Discovery, Intelligent Systems, and Distributed Analytics* **2024**, *14*, 1–13.

Chung, J.; Ko, R.; Yoo, W.; Onizuka, M.; Kim, S.; Kim, T.W.; Shin, W.Y. GraphCompliance: Aligning Policy and Context Graphs for LLM-Based Regulatory Compliance. *arXiv preprint arXiv:2510.26309* **2025**.

 doi:10.20944/preprints202602.1970.v2

Tan, Y.; Wu, B.; Cao, J.; Jiang, B. LLaMA-UTP: Knowledge-Guided Expert Mixture for Analyzing Uncertain Tax Positions. *IEEE Access* **2025**.

Hernandez, J.; Golpayegani, D.; Lewis, D. An open knowledge graph-based approach for mapping concepts and requirements between the eu ai act and international standards. *AI and Ethics* **2025**, pp. 1–12.

Usmanova, A.; Usbeck, R. Structuring sustainability reports for environmental standards with LLMs guided by ontology. In Proceedings of the Proceedings of the 1st Workshop on Natural Language Processing Meets Climate Change (ClimateNLP 2024), 2024, pp. 168–177.

Ong, K.; Mao, R.; Varshney, D.; Xing, F.; Satapathy, R.; Sulaeman, J.; Cambria, E.; Mengaldo, G. ESGSenticNet: A neurosymbolic knowledge base for corporate sustainability analysis. *arXiv preprint arXiv:2501.15720* **2025**.

Osman, C.C.; Ghiran, A.M.; Buchmann, R.A. Towards a Knowledge Management Capability for ESG Accounting with the Help of Enterprise Modeling and Knowledge Graphs. In Proceedings of the Companion Proceedings of the 17th IFIP WG 8.1 Working Conference on the Practice of Enterprise Modeling Forum, M4S, FACETE, AEM, Tools and Demos co-located with PoEM 2024, 2024.

Dolha, D.N.; Osman, C.C.; Chiș, A. KM4ESG: BPMN and AI-powered knowledge management platform for ESG analysis and reporting **2025**.

Wu, W.; Wen, C.; Yuan, Q.; Chen, Q.; Cao, Y. Construction and application of knowledge graph for construction accidents based on deep learning. *Engineering, construction and architectural management* **2025**, *32*, 1097–1121.

Liu, Q.; Ding, Y.; Luo, X. Automated knowledge graph-based risk assessment for fall-from-height accidents in construction. *Automation in Construction* **2025**, *179*, 106482.

Chen, Q.; Long, D.; Yang, C.; Xu, H. Knowledge graph improved dynamic risk analysis method for behavior-based safety management on a construction site. *Journal of Management in Engineering* **2023**, *39*, 04023023.

Shi, F.; Wu, Z. Construction of knowledge graph of the elevator safety accidents and analysis of key risk factors based on KG-DEMATEL-ISM-MICMAC method. *IEEE Access* **2024**, *12*, 43615–43631.

Zhang, Z.; Wang, J.; Li, B.; Lin, X.; Liu, M. Construction of a Person–Job Temporal Knowledge Graph Using Large Language Models. *Big Data and Cognitive Computing* **2025**, *9*, 287.

Fettach, Y.; Bahaj, A.; Ghogho, M. Skill Demand Forecasting Using Temporal Knowledge Graph Embeddings. *arXiv preprint arXiv:2504.07233* **2025**.

Liu, H.; Zhang, N.; Su, Z. A Knowledge Graph Framework for Purpose-Driven Privacy Policy Compliance Auditing. In Proceedings of the International Symposium on Knowledge and Systems Sciences. Springer, 2025, pp. 69–83.

Cheng, Y.; Bajaber, O.; Tsegai, S.A.; Song, D.; Gao, P. Ctinexus: Automatic cyber threat intelligence knowledge graph construction using large language models. In Proceedings of the 2025 IEEE 10th European Symposium on Security and Privacy (EuroS&P). IEEE, 2025, pp. 923–938.

Hu, Y.; Zou, F.; Han, J.; Sun, X.; Wang, Y. Llm-tikg: Threat intelligence knowledge graph construction utilizing large language model. *Computers & Security* **2024**, *145*, 103999.

Cotti, L.; Drago, I.; Rula, A.; Bianchini, D.; Cerutti, F. OntoLogX: Ontology-Guided Knowledge Graph Extraction from Cybersecurity Logs with Large Language Models. *arXiv preprint arXiv:2510.01409* **2025**.

Kurniawan, K.; Kiesling, E.; Ekelhart, A. CyKG-RAG: Towards knowledge-graph enhanced retrieval augmented generation for cybersecurity **2024**.

Angioni, S.; Consoli, S.; Dessì, D.; Osborne, F.; Recupero, D.R.; Salatino, A. Exploring environmental, social, and governance (esg) discourse in news: An ai-powered investigation through knowledge graph analysis. *IEEE Access* **2024**, *12*, 77269–77283.

Hassan Nassar, O.M.; Jafari, F.; Jain, C. From News to Knowledge: Leveraging AI and Knowledge Graphs for Real-Time ESG Insights. *Sustainability* **2025**, *17*, 11128.

Iwata, T.; Comte, G.; Flores, M.; Kondo, R.; Hisano, R. Aligning ESG Controversy Data with International Guidelines through Semi-Automatic Ontology Construction. *arXiv preprint arXiv:2509.10922* **2025**.

Islam, M.S. KnowUREnvironment: An Automated Knowledge Graph for Climate Change and Environmental Issues. In Proceedings of the AAAI 2022 Fall Symposium: The Role of AI in Responding to Climate Challenges, 2022.

Penev, L.; Koureas, D.; Groom, Q.; Lanfear, J.; Agosti, D.; Casino, A.; Miller, J.; Arvanitidis, C.; Cochrane, G.; Hobern, D.; et al. Biodiversity community integrated knowledge library (BiCIKL). *Research Ideas and Outcomes* **2022**, *8*, e81136.

Zhao, L.; Sun, Y.; Ren, J.; Gao, H.; Xiao, G. Construction of waste-to-resource knowledge graph for industrial symbiosis identification using large language models. *Nature Communications* **2025**.

 doi:10.20944/preprints202602.1970.v2

Zhang, Y.; Du, T.; Ma, Y.; Wang, X.; Xie, Y.; Yang, G.; Lu, Y.; Chang, E.C. Attackg+: Boosting attack knowledge graph construction with large language models. *arXiv preprint arXiv:2405.04753* **2024**.

Ren, Y.; Xiao, Y.; Zhou, Y.; Zhang, Z.; Tian, Z. CSKG4APT: A cybersecurity knowledge graph for advanced persistent threat organization attribution. *IEEE Transactions on Knowledge and Data Engineering* **2022**, *35*, 5695–5709.

Zhu, R.; Shimizu, C.; Stephen, S.; Fisher, C.K.; Thelen, T.; Currier, K.; Janowicz, K.; Hitzler, P.; Schildhauer, M.; Li, W.; et al. The knowwheregraph: A large-scale geo-knowledge graph for interdisciplinary knowledge discovery and geo-enrichment. *arXiv preprint arXiv:2502.13874* **2025**.

van der Weerdt, R.; de Boer, V.; Siebes, R.; Groenewold, R.; van Harmelen, F. OfficeGraph: a knowledge graph of office building IoT measurements. In Proceedings of the European Semantic Web Conference. Springer, 2024, pp. 94–109.

Wu, Z.; Islam, S.; Tang, L. Visualizing ESG Performance in an Integrated GIS–BIM–IoT Platform for Strategic Urban Planning. *Buildings* **2025**, *15*, 3394.

Keena, N.; Friedman, A.; Parsaee, M.; Mussio, M.; Klein, A.; Pomasonco-Alvis, M.; Pinheiro, P. Housing Passport knowledge graph: Promoting a circular economy in urban residential buildings. *Sustainable Cities and Society* **2025**, *119*, 106050.

Jiang, W.; Liu, Y.; Chen, K.; Liu, Y.; Ding, L. Early-warning of unsafe hoisting operations: An integration of digital twin and knowledge graph. *Developments in the Built Environment* **2024**, *19*, 100490.

Wen, H. A Model of Proactive Safety Based on Knowledge Graph. *arXiv preprint arXiv:2407.15127* **2024**.

Ren, S.; Ma, Z.; Wang, M. Product Qulity Traceability of Cold Chain Logistics Based on Multimodal Knowledge Graph. In Proceedings of the 2024 20th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD). IEEE, 2024, pp. 1–6.

Wang, S.; Yang, J.; Yang, B.; Li, D.; Kang, L. An intelligent quality control method for manufacturing processes based on a Human–Cyber–Physical knowledge graph. *Engineering* **2024**, *41*, 242–260.

Aivalis, T.; Klampanos, I.A.; Troumpoukis, A. LLM-Driven Knowledge Graph Construction from Earth Observation Data for Extreme Events **2025**.

Chen, Y.; Zhang, L.; Chen, L.; Shi, Y. Construction of Knowledge Graphs for the Constituent Elements and Mineralization Process of Urban Minerals: A Case of Iron and Steel Resources. *Sustainability* **2025**, *17*, 4136.

Quek, H.Y.; Hofmeister, M.; Rihm, S.D.; Yan, J.; Lai, J.; Brownbridge, G.; Hillman, M.; Mosbach, S.; Ang, W.; Tsai, Y.K.; et al. Dynamic knowledge graph applications for augmented built environments through "The World Avatar". *Journal of Building Engineering* **2024**, *91*, 109507.

Karjou, P.F.; Berktold, M.; Saryazdi, S.K.; Rätz, M.; Müller, D. Integrating IoT, large language models, and knowledge graphs for future smart districts: A semantic approach for energy performance assessment. In Proceedings of the Journal of Physics: Conference Series. IOP Publishing, 2025, Vol. 3140, p. 042006.

Saad, M.; Zhang, Y.; Tian, J.; Jia, J. A graph database for life cycle inventory using Neo4j. *Journal of Cleaner Production* **2023**, *393*, 136344.

Wu, J.; Orlandi, F.; O'Sullivan, D.; Dev, S. LinkClimate: An interoperable knowledge graph platform for climate data. *Computers & Geosciences* **2022**, *169*, 105215.

Delgoshaei, P.; Heidarinejad, M.; Austin, M.A. A semantic approach for building system operations: Knowledge representation and reasoning. *Sustainability* **2022**, *14*, 5810.

Liu, C.; Yang, S. Using text mining to establish knowledge graph from accident/incident reports in risk assessment. *Expert Systems with Applications* **2022**, *207*, 117991.

Liu, D.; Cheng, L. MAKG: A maritime accident knowledge graph for intelligent accident analysis and management. *Ocean Engineering* **2024**, *312*, 119280.

Li, Y.; Xu, F.; Zhang, Z.; Mei, X. Construction Behavior Safety Risk Identification Based on Knowledge Graph and Large Language Model. In *Computing in Civil Engineering 2024*; pp. 895–903.

Shen, G.; Wang, W.; Mu, Q.; Pu, Y.; Qin, Y.; Yu, M. Data-Driven Cybersecurity Knowledge Graph Construction for Industrial Control System Security. *Wireless Communications and Mobile Computing* **2020**, *2020*, 8883696.

Yin, J.; Hong, W.; Wang, H.; Cao, J.; Miao, Y.; Zhang, Y. A compact vulnerability knowledge graph for risk assessment. *ACM Transactions on Knowledge Discovery from Data* **2024**, *18*, 1–17.

Oksanen, J.; Majumder, A.; Saunack, K.; Toni, F.; Dhondiyal, A. A graph-based method for unsupervised knowledge discovery from financial texts. In Proceedings of the Proceedings of the Thirteenth Language Resources and Evaluation Conference, 2022, pp. 5412–5417.

Shen, Y.; Guo, C.; Li, H.; Chen, J.; Guo, Y.; Qiu, X. Financial feature embedding with knowledge representation learning for financial statement fraud detection. *Procedia Computer Science* **2021**, *187*, 420–425.

 doi:10.20944/preprints202602.1970.v2

Wu, H.; Chang, Y.; Li, J.; Zhu, X. Financial fraud risk analysis based on audit information knowledge graph. *Procedia Computer Science* **2022**, *199*, 780–787.

Zhang, J.; Zhang, W.; Wei, Q.; Zhong, H.; Miao, Y. GRAG-ProSafe QAS: Graph retrieval-augmented generation for process production safety intelligent question and answer system. *Expert Systems with Applications* **2025**, p. 129947.

Jomraj, H.S.; Agarwal, B.; Rojkova, V. RAGulating Compliance: A Multi-Agent Schema-Light Knowledge Graph for Regulatory Compliance QA **2025**.

Felder, M.; De Marchi, M.; Dallasega, P.; Rauch, E. Smart Routing for Sustainable Supply Chain Networks: An AI and Knowledge Graph Driven Approach. *Applied Sciences* **2025**, *15*, 8001.

Hou, Y.; Qiao, D.; Han, H.; Huang, J. Knowledge Graph-Based Dynamic Differential Evolution Algorithm for E-Waste Recycling Vehicle Routing Problem. In Proceedings of the 2025 IEEE 6th International Seminar on Artificial Intelligence, Networking and Information Technology (AINIT). IEEE, 2025, pp. 1–6.

Yan, J.; Lv, T.; Yu, Y. Construction and recommendation of a water affair knowledge graph. *Sustainability* **2018**, *10*, 3429.

Elluri, L.; Nagar, A.; Joshi, K.P. An integrated knowledge graph to automate gdpr and pci dss compliance. In Proceedings of the 2018 IEEE International Conference on Big Data (Big Data). IEEE, 2018, pp. 1266–1271.

Atzeni, P.; Bellomarini, L.; Iezzi, M.; Sallinger, E.; Vlad, A. Weaving Enterprise Knowledge Graphs: The Case of Company Ownership Graphs. In Proceedings of the EDBT, 2020, pp. 555–566.

Mishra, P.; Mittal, R. NeuralNERE: Neural Named Entity Relationship Extraction for End-to-End Climate Change Knowledge Graph Construction. In Proceedings of the ICML 2021 Workshop on Tackling Climate Change with Machine Learning, 2021.

Vanapalli, K.; Kilaru, A.; Shafiq, O.; Khan, S. Unifying Large Language Models and Knowledge Graphs for efficient Regulatory Information Retrieval and Answer Generation. *COLING 2025* **2025**, p. 22.

Ushio, K.; Tsuji, D.; Kobashi, Y. GraphRAG with Knowledge Graphs for Question Answering on Administrative Meeting Records **2025**.

DeBellis, M.; Gino, G.; Balaji, A.; Gino, J. Using Retrieval Augmented Generation and Knowledge Graphs to Understand Climate Obstruction **2025**.

Kaoukis, G.; Koufopoulos, I.A.; Psaroudaki, E.; Karidi, D.P.; Pitoura, E.; Papastefanatos, G.; Tsaparas, P. Emerald-Mind: A Knowledge Graph-Augmented Framework for Greenwashing Detection. 2025.

Wang, Y.; Huang, T.Y.; Gao, Q.; Zhang, J. HeDA: An Intelligent Agent System for Heatwave Risk Discovery through Automated Knowledge Graph Construction and Multi-layer Risk Propagation Analysis. *arXiv preprint arXiv:2509.25112* **2025**.

Cai, S.; Xie, Z. Explainable fraud detection of financial statement data driven by two-layer knowledge graph. *Expert Systems with Applications* **2024**, *246*, 123126.

Wang, S.; Zhang, Z.; Fang, L.; Nguyen, C.T.; Li, W. Corporate Fraud Detection in Rich-yet-Noisy Financial Graph. *arXiv preprint arXiv:2502.19305* **2025**.

Zhu, S.; Ma, T.; Wu, H.; Ren, J.; He, D.; Li, Y.; Ge, R. Expanding and Interpreting Financial Statement Fraud Detection Using Supply Chain Knowledge Graphs. *Journal of Theoretical and Applied Electronic Commerce Research* **2025**, *20*, 26.

Benjira, W.; Travers, N.; Atigui, F.; Bucher, B.; Grim-Yefsah, M. SDG-KG: A Framework to Compute SDG Indicators with Open Data. *Proceedings of the VLDB Endowment* **2025**, *18*, 5367–5370.

Zhao, Y.; Dai, C.; Niyato, D.; Tan, C.F.; Xiang, K.; Wang, Y.; Yeo, Z.; Loong, D.T.Z.; Zhaozhi, J.L.; HO, E.H. A Graph-Retrieval-Augmented Generation Framework Enhances Decision-Making in the Circular Economy. *arXiv preprint arXiv:2506.04252* **2025**.

Lu, Z.; Sun, C.; Hu, Y.; Kumar, A. BIM and Knowledge Graph-Based Building Material Recycle and Reuse Assessment Framework. In *Computing in Civil Engineering 2023*; 2024; pp. 517–525.

Yin, L.; Wang, C.; Cong, L.; Du, Q. A sequential disassembly planning approach based on knowledge graph and graph isomorphism network for supporting power battery remanufacturing. *Journal of Cleaner Production* **2025**, p. 145558.

Tauqeer, A.; Kurteva, A.; Chhetri, T.R.; Ahmeti, A.; Fensel, A. Automated GDPR contract compliance verification using knowledge graphs. *Information* **2022**, *13*, 447.

Zhang, T. A Knowledge Graph-Enhanced Multimodal AI Framework for Intelligent Tax Data Integration and Compliance Enhancement. *Frontiers in Business and Finance* **2025**, *2*, 247–261.

 doi:10.20944/preprints202602.1970.v2

Kosasih, E.E.; Margaroli, F.; Gelli, S.; Aziz, A.; Wildgoose, N.; Brintrup, A. Towards knowledge graph reasoning for supply chain risk management using graph neural networks. *International Journal of Production Research* **2024**, *62*, 5596–5612.

AlMahri, S.; Xu, L.; Brintrup, A. Enhancing supply chain visibility with knowledge graphs and large language models. *International Journal of Production Research* **2025**, pp. 1–32.

Jin, B.; Sun, Q.; Chen, L. Enhancing Supply Chain Transparency in Emerging Economies Using Online Contents and LLMs. In Proceedings of the 2025 International Conference on Information Networking (ICOIN). IEEE, 2025, pp. 487–492.

Heus, E.; Bookstaber, R.; Sharma, D. Exploring Network-Knowledge Graph Duality: A Case Study in Agentic Supply Chain Risk Analysis. *arXiv preprint arXiv:2510.01115* **2025**.

Zheng, S.; Mizushina, K.; Naono, K. Empowering Supply Chain Risk Monitoring with Ontology-Guided Knowledge Graph Extraction by LLMs **2025**.

Marotta, S.M.; Masucci, V.; Kontogiannis, S.; Avgerinakis, K. From Unified Ontology to Knowledge Base: Data Fusion for Enhanced Wildfire Management. In Proceedings of the International conference on WorldS4. Springer, 2024, pp. 131–150.

Burel, G.; Alani, H. ClimaFactsKG: Towards an Interlinked Knowledge Graph of Scientific Evidence to Fight Climate Misinformation. In Proceedings of the CEUR Workshop Proceedings, 2025, Vol. 4065, pp. 134–140.

Papageorgiou, N.; Pournara, D.; Apostolou, D.; Mentzas, G. A Graph-Based Approach for Representing Water Treatment Process Knowledge. In Proceedings of the 2023 14th International Conference on Information, Intelligence, Systems & Applications (IISA). IEEE, 2023, pp. 1–8.

Papageorgiou, N.; Pournara, D.; Apostolou, D.; Mentzas, G. Managing industrial water treatment processes knowledge with knowledge graphs. *Intelligent Decision Technologies* **2025**, *19*, 687–717.

Jing, R.; Li, P. Knowledge graph for integration and quality traceability of agricultural product information. *Frontiers in sustainable food systems* **2024**, *8*, 1389945.

Zang, T.; Yang, M.; Jiang, P. Event State Knowledge Graph-driven Full-chain Product Quality Tracing. *Procedia CIRP* **2025**, *134*, 538–543.

Wu, H.; Jiang, Z.; Zhu, S.; Zhang, H. A knowledge graph based disassembly sequence planning for end-of-life power battery. *International Journal of Precision Engineering and Manufacturing-Green Technology* **2024**, *11*, 849–861.

Wang, L.; Liu, X.; Liu, Y.; Li, H.; Liu, J.; Yang, L. Knowledge graph-based method for intelligent generation of emergency plans for water conservancy projects. *IEEE Access* **2023**, *11*, 84414–84429.

Fu, C.; Huang, Z.; van Harmelen, F.; He, T.; Jiang, X. Food4healthKG: Knowledge graphs for food recommendations based on gut microbiota and mental health. *Artificial Intelligence in Medicine* **2023**, *145*, 102677.

Gao, F.; Zhao, X.; Xia, D.; Zhou, Z.; Yang, R.; Lu, J.; Jiang, H.; Park, C.; Li, I. HealthGenie: A Knowledge-Driven LLM Framework for Tailored Dietary Guidance. In Proceedings of the Proceedings of the 34th ACM International Conference on Information and Knowledge Management, 2025, pp. 6639–6643.

Weichselbraun, A.; Waldvogel, R.; Fraefel, A.; van Schie, A.; Kuntschik, P. Building knowledge graphs and recommender systems for suggesting reskilling and upskilling options from the web. *Information* **2022**, *13*, 510.

Fathi, F. AIREG: Enhanced Educational Recommender System with Large Language Models and Knowledge Graphs. In Proceedings of the European Semantic Web Conference. Springer, 2024, pp. 84–93.

Cai, B.; Ye, Z.; Chen, S. Intelligent ESG Evaluation for Construction Enterprises in China: An LLM-Based Model. *Buildings* **2025**, *15*, 2710.

Fotopoulou, E.; Mandilara, I.; Zafeiropoulos, A.; Laspidou, C.; Adamos, G.; Koundouri, P.; Papavassiliou, S. SustainGraph: A knowledge graph for tracking the progress and the interlinking among the sustainable development goals' targets. *Frontiers in Environmental Science* **2022**, *10*, 1003599.

Androna, C.M.; Mandilara, I.; Fotopoulou, E.; Zafeiropoulos, A.; Papavassiliou, S. A Knowledge Graph-Driven Analysis of the Interlinkages among the Sustainable Development Goal Indicators in Different Spatial Resolutions. *Sustainability* **2024**, *16*, 4328.

He, C.; Zhou, X.; Wu, Y.; Yu, X.; Zhang, Y.; Zhang, L.; Wang, D.; Lyu, S.; Xu, H.; Xiaoqiao, W.; et al. Esgenius: Benchmarking llms on environmental, social, and governance (esg) and sustainability knowledge. In Proceedings of the Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing, 2025, pp. 14623–14664.

 doi:10.20944/preprints202602.1970.v2

Zhang, L.; Zhou, X.; He, C.; Wang, D.; Wu, Y.; Xu, H.; Liu, W.; Miao, C. Mmesgbench: Pioneering multimodal understanding and complex reasoning benchmark for esg tasks. In Proceedings of the Proceedings of the 33rd ACM International Conference on Multimedia, 2025, pp. 12829–12836.

Hogan, A.; Blomqvist, E.; Cochez, M.; d'Amato, C.; Melo, G.D.; Gutierrez, C.; Kirrane, S.; Gayo, J.E.L.; Navigli, R.; Neumaier, S.; et al. Knowledge graphs. *ACM Computing Surveys (Csur)* **2021**, *54*, 1–37.

Ji, S.; Pan, S.; Cambria, E.; Marttinen, P.; Yu, P.S. A survey on knowledge graphs: Representation, acquisition, and applications. *IEEE transactions on neural networks and learning systems* **2021**, *33*, 494–514.

Zhu, Y.; Wang, X.; Chen, J.; Qiao, S.; Ou, Y.; Yao, Y.; Deng, S.; Chen, H.; Zhang, N. Llms for knowledge graph construction and reasoning: Recent capabilities and future opportunities. *World Wide Web* **2024**, *27*, 58.

Peng, B.; Zhu, Y.; Liu, Y.; Bo, X.; Shi, H.; Hong, C.; Zhang, Y.; Tang, S. Graph retrieval-augmented generation: A survey. *ACM Transactions on Information Systems* **2024**.

He, C.; Zhou, X.; Yu, X.; Zhang, L.; Zhang, Y.; Wu, Y.; Xiao, L.; Li, L.; Wang, D.; Xu, H.; et al. SSKG Hub: An Expert-Guided Platform for LLM-Empowered Sustainability Standards Knowledge Graphs. *arXiv preprint arXiv:2603.00669* **2026**.

He, C.; Zhou, X.; Wang, D.; Xu, H.; Liu, W.; Miao, C. PCA-OS: A Planetary Climate Adaptation Operating System **2026**.

Gusenbauer, M. Google Scholar to overshadow them all? Comparing the sizes of 12 academic search engines and bibliographic databases. *Scientometrics* **2019**, *118*, 177–214.

Bramer, W.M.; De Jonge, G.B.; Rethlefsen, M.L.; Mast, F.; Kleijnen, J. A systematic approach to searching: an efficient and complete method to develop literature searches. *Journal of the Medical Library Association: JMLA* **2018**, *106*, 531.

Moher, D.; Liberati, A.; Tetzlaff, J.; Altman, D.G.; Group, P.; et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *International journal of surgery* **2010**, *8*, 336–341.

Fu, Y.; Wen, P.; Wu, J.; Shu, Y. Exploring knowledge structure of "dual carbon" policies: Combining computational text mining and knowledge graph. *Energy Strategy Reviews* **2025**, *62*, 101976.

Ji, X.; Fu, Q.; Wang, H.; Gao, Z. Carbon Emission Reduction Path Analysis and Decision Support Based on Knowledge Graphs. In Proceedings of the 2025 2nd International Symposium on New Energy Technologies and Power Systems (NETPS). IEEE, 2025, pp. 363–369.

Sheng, X.; Shi, W.; Wei, S.; Yang, Q.; Wang, S.; Li, M. Construction of a Spatiotemporal Knowledge Graph for Carbon Source and Sink Dynamics Driven by Multi-Source Sensing Collaboration. In Proceedings of the 2025 32nd International Conference on Geoinformatics. IEEE, 2025, pp. 1–6.

Xie, W.; Farazi, F.; Atherton, J.; Bai, J.; Mosbach, S.; Akroyd, J.; Kraft, M. Dynamic knowledge graph approach for modelling the decarbonisation of power systems. *Energy and AI* **2024**, *17*, 100359.

Chatterjee, A.; Ranganathan, S. Representation Learning of Complex Assemblies, An Effort to Improve Corporate Scope 3 Emissions Calculation. *arXiv preprint arXiv:2409.03769* **2024**.

Zhao, Z. Quantification of Carbon Emission Technologies Based on Knowledge Graph Bert-BiLSTM-Attention-CRF Model. In Proceedings of the 2023 International Conference on Electronics and Devices, Computational Science (ICEDCS). IEEE, 2023, pp. 42–47.

Wu, J.; Pierse, J.; Orlandi, F.; O'Sullivan, D.; Dev, S. Improving tourism analytics from climate data using knowledge graphs. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **2023**, *16*, 2402–2412.

Oladeji, O.; Mousavi, S.S. Leveraging AI-Derived Data for Carbon Accounting: Information Extraction from Alternative Sources. In Proceedings of the Proceedings of the AAAI Symposium Series, 2023, Vol. 2, pp. 135–139.

Ma, Z.; Zhang, S.; Jia, H.; Liu, K.; Xie, X.; Qu, Y. A knowledge graph-based approach to recommending low-carbon construction schemes of bridges. *Buildings* **2023**, *13*, 1352.

Oladeji, O.; Mousavi, S.S.; Roston, M. AI-driven E-Liability Knowledge Graphs: A Comprehensive Framework for Supply Chain Carbon Accounting and Emissions Liability Management. In Proceedings of the Proceedings of the AAAI Symposium Series, 2023, Vol. 2, pp. 124–134.

Sharma, S.; Roy Chowdhury, M.; Sirmokadam, S. Carbon footprint analysis using knowledge graph. In *Intelligent Sustainable Systems: Selected Papers of WorldS4 2021, Volume 1*; Springer, 2022; pp. 587–595.

Stade, C.; Schneider, J.; Fu, Y. From Evidence to Insights: GraphRAG as a Dynamic Knowledge Layer for the Collaboration for Environmental Evidence's Database of Evidence Reviews **2025**.

Wu, Z.; Yang, H.; Cai, Y.; Yu, B.; Liang, C.; Duan, Z.; Liang, Q. Intelligent monitoring applications of landslide disaster knowledge graphs based on ChatGLM2. *Remote Sensing* **2024**, *16*, 4056.

 doi:10.20944/preprints202602.1970.v2

Androna, C.M.; Mandilara, I.; Fotopoulou, E.; Zafeiropoulos, A.; Papavassiliou, S.; Ziliaskopoulos, K.; Laspidou, C. An open and interoperable Knowledge Management framework for Climate Change Vulnerability Assessment. *ACM Journal on Computing and Sustainable Societies* **2024**, *2*, 1–30.

Tu, S.; Zhuang, W.; Ren, F. Constructing a Knowledge Graph for Extreme Climate Architecture Based on Large Language Models (LLMs). In Proceedings of the The International Conference on Computational Design and Robotic Fabrication. Springer, 2024, pp. 251–261.

Simsek-Senel, G.; Rijgersberg, H.; Öztürk, B.; Weits, J.; Fensel, A. I-know-foo: interlinking and creating knowledge graphs for near-zero co2 emission diets and sustainable food production. In Proceedings of the International Symposium on AI, Data and Digitalization. Springer Nature Switzerland Cham, 2023, pp. 106–119.

Sheikh, H.A.; Kushwaha, N.; Singh, A. NATUREKG: AN ONTOLOGY AND KNOWLEDGE GRAPH FOR NATURE FINANCE WITH A TEXT2CYPHER APPLICATION. *Frontiers in Artificial Intelligence*, *8*, 1693843.

Vasiliu, L.; Dizaji, S.; Eberhart, A.; Roman, D.; Prodan, R. Modeling and generating extreme volumes of financial synthetic time-series data with knowledge graphs. In Proceedings of the Proceedings of the 26th CEUR Workshop on Financial Knowledge Graphs, 2024.

Peng, Y.; Guo, J.; Liu, Y. Comparative Analysis of China's Pilot Carbon Market Based on Knowledge Graph. In Proceedings of the 2023 3rd Asia-Pacific Conference on Communications Technology and Computer Science (ACCTCS). IEEE, 2023, pp. 521–525.

Zhoua, X.; Zhanga, L.; Leungb, C. Green Premium: Evaluating and Diminishing the Environmental Surcharge. *International Journal of Information Technology* **2022**, *28*.

Wu, P.; Tu, H.; Mou, X.; Gong, L. An intelligent energy management method for the manufacturing systems using the knowledge graph and large language model. *Journal of Intelligent Manufacturing* **2025**, pp. 1–20.

Hanžel, V.; Bertalanič, B.; Fortuna, C. Towards data-driven electricity management: multi-region uniform data and knowledge graph. *Scientific Data* **2025**, *12*, 38.

Wang, R.G.; Ho, W.J.; Chiang, K.C.; Hung, Y.C.; Tai, J.K.; Tan, J.C.; Chuang, M.L.; Ke, C.Y.; Chien, Y.F.; Jeng, A.P.; et al. Analyzing long-term and high instantaneous power consumption of buildings from smart meter big data with deep learning and knowledge graph techniques. *Energies* **2023**, *16*, 6893.

Janev, V.; Vidal, M.E.; Pujić, D.; Popadić, D.; Iglesias, E.; Sakor, A.; Čampa, A. Responsible knowledge management in energy data ecosystems. *Energies* **2022**, *15*, 3973.

Li, T.; Zhao, Y.; Zhang, C.; Zhou, K.; Zhang, X. A semantic model-based fault detection approach for building energy systems. *Building and Environment* **2022**, *207*, 108548.

Aryan, P.R.; Ekaputra, F.J.; Sabou, M.; Hauer, D.; Mosshammer, R.; Einfalt, A.; Miksa, T.; Rauber, A. Explainable cyber-physical energy systems based on knowledge graph. In Proceedings of the Proceedings of the 9th Workshop on Modeling and Simulation of Cyber-Physical Energy Systems, 2021, pp. 1–6.

Xiaoping, G.; Mengyu, R.; Hong, Z.; Ping, W.; Ruijun, R.; Feng, G. Construction technology of knowledge graph and its application in power grid. In Proceedings of the E3S Web of Conferences. EDP Sciences, 2021, Vol. 256, p. 01039.

Janev, V.; Vidal, M.E.; Endris, K.; Pujic, D. Managing knowledge in energy data spaces. In Proceedings of the Companion Proceedings of the Web Conference 2021, 2021, pp. 7–15.

Wang, C.; An, J.; Mu, G. Power system network topology identification based on knowledge graph and graph neural network. *Frontiers in Energy Research* **2021**, *8*, 613331.

Jiang, G.; Su, L.; Liu, H.; Cao, Y.; Sun, R.; Diao, F. Constructing the power knowledge graph by multi-source electricity data. In Proceedings of the 2020 International Conference on Computer, Information and Telecommunication Systems (CITS). IEEE, 2020, pp. 1–5.

Zhao, L.; Zhao, Z.; Xu, H.; Zhang, Y.; Xu, Y. Visual analysis and mining of knowledge graph for power network data based on natural language processing. In Proceedings of the 2020 Fourth International Conference on Computing Methodologies and Communication (ICCMC). IEEE, 2020, pp. 410–413.

Tang, Y.; Liu, T.; Liu, G.; Li, J.; Dai, R.; Yuan, C. Enhancement of power equipment management using knowledge graph. In Proceedings of the 2019 IEEE Innovative Smart Grid Technologies-Asia (ISGT Asia). IEEE, 2019, pp. 905–910.

Guo, D.; Di Modica, P.; La Rosa, A.D. A Semantic-enriched LCI database for the Embodied Environmental Evaluation of Buildings through Knowledge Graph Technologies. *IFAC-PapersOnLine* **2025**, *59*, 2474–2478.

Yan, R.; Chen, Z.; Zhou, S.; Niu, G.; Li, Y.; Liu, Z.; Wang, J.; Wu, X.; Luo, Q.; Zhou, Y.; et al. ForestFoodKG: A Structured Dataset and Knowledge Graph for Forest Food Taxonomy and Nutrition. *Foods* **2025**, *14*, 4186.

Shaw, C.; Hoare, C.; de Riet, M.; de Andrade Pereira, F.; O'Donnell, J. An end-to-end Asset Life Cycle Knowledge Graph. In Proceedings of the Proceedings of the 2024 W78 Conference, CIB W, Vol. 78, pp. 81–90.

 doi:10.20944/preprints202602.1970.v2

Greif, L.; Hauck, S.; Kimmig, A.; Ovtcharova, J. A Knowledge Graph Framework to Support Life Cycle Assessment for Sustainable Decision-Making. *Applied Sciences* **2024**, *15*, 175.

Peng, T.; Gao, L.; Agbozo, R.S.; Xu, Y.; Svynarenko, K.; Wu, Q.; Li, C.; Tang, R. Knowledge graph-based mapping and recommendation to automate life cycle assessment. *Advanced Engineering Informatics* **2024**, *62*, 102752.

Li, J.; Li, S.; Jiang, J.; Zhang, F.; Liu, C. A Method for Generating Carbon Footprint Inventory of Substation Engineering Based on Knowledge Graph. In Proceedings of the 2024 IEEE 2nd International Conference on Image Processing and Computer Applications (ICIPCA). IEEE, 2024, pp. 88–92.

Wang, Y.; Tao, J.; Liu, W.; Peng, T.; Tang, R.; Wu, Q. A knowledge-enriched framework for life cycle assessment in manufacturing. *Procedia CIRP* **2022**, *105*, 55–60.

Yan, R.; An, P.; Meng, X.; Li, Y.; Li, D.; Xu, F.; Dang, D. A knowledge graph for crop diseases and pests in China. *Scientific Data* **2025**, *12*, 222.

Tabanao, G.; Pagdanganan, A.M.; Batista-Navarro, R.; Gabud, R. Towards the Curation of Environment-related Knowledge Graphs: Fine-tuning General-domain Language Models for Biodiversity Named Entity Recognition. In Proceedings of the ICLR 2025 Workshop on Tackling Climate Change with Machine Learning, 2025.

Huseynov, J.; Dalal, S.; Shepelenko, V. Firmographica: Knowledge Graph and AI-based Framework for Short-Selling Risk Assessment. *Journal of Computer Science and Digital Technologies* **2025**, *1*, 39–49.

Hamed, N.; Rana, O.; Orozco Ter Wengel, P.; Goossens, B.; Perera, C. FOODS: Ontology-based knowledge graphs for forest observatories. *ACM Journal on Computing and Sustainable Societies* **2025**, *3*, 1–42.

Xiao, X.; Wang, P.; Ge, Y.; Luo, J.; Chen, H.; He, Y.; Zhang, D.; Li, Y.; Fang, C.; Lin, H. GeoKG-HSA: A framework for habitat suitability assessment with geospatial knowledge graphs. *International Journal of Applied Earth Observation and Geoinformation* **2025**, *144*, 104921.

Shirvani-Mahdavi, N.; Wingfield, D.; Gutierrez, J.G.; Tran, M.; Zhu, Z.; Zhang, Z.; Zhang, H.; Goudar, A.D.; Li, C.; Jin, V.; et al. A knowledge graph informing soil carbon modeling. In Proceedings of the International Conference on Web Engineering. Springer, 2025, pp. 226–241.

Ahmeti, A.; Hensel, D.S.; Pruski, C.; Tyc, J.; Hensel, M. Enabling Biodiversity-Informed Architecture Through Ontology-Driven Data Integration. *Applied Sciences* **2025**, *15*, 5311.

Tandon, D.; De Farias, T.M.; Allard, P.M.; Defossez, E. METRIN-KG: A knowledge graph integrating plant metabolites, traits and biotic interactions. *bioRxiv* **2025**, pp. 2025–08.

Martinez Otalora, J.D.; Shen, J.; Rojas Celis, A.P. Knowledge graph-enhanced pattern language for biodiversity integration in architectural education. *Humanities and Social Sciences Communications* **2025**, *12*, 1762.

Ashworth, A.J.; Tyson, A.; Propst, T.; Marshall, L.; Li, C.; Volenec, J.J.; Berti, M.T.; Picasso, V.; Foster, J.L.; Su, J. Knowledge graph applications for identifying resilient forage systems. *Agricultural & Environmental Letters* **2025**, *10*, e70021.

Bartalesi, V.; Coro, G.; Lenzi, E.; Pratelli, N.; Pagano, P.; Moretti, M.; Brunori, G. A Semantic Knowledge Graph of European Mountain Value Chains. *Scientific Data* **2024**, *11*, 978.

Wen, X.; Shi, J.; Chen, S. Visualizing and analyzing the impact of Chinese forests on the environment using a knowledge graph. In Proceedings of the Proceedings of the 2024 8th International Conference on Electronic Information Technology and Computer Engineering, 2024, pp. 60–64.

Ahmeti, A.; Schakel, J.K.; David, R.; Revenko, A. Towards preserving Biodiversity using Nature FIRST Knowledge Graph with Crossovers. In Proceedings of the ISWC (Posters/Demos/Industry), 2023.

Le Guillarme, N.; Thuiller, W. A practical approach to constructing a knowledge graph for soil ecological research. *European Journal of Soil Biology* **2023**, *117*, 103497.

Escamilla Molgora, J.M.; Sedda, L.; Atkinson, P.M. Biospytial: spatial graph-based computing for ecological Big Data. *GigaScience* **2020**, *9*, giaa039.

Liu, Y.; He, B.; Hildebrandt, M.; Buchner, M.; Inzko, D.; Wernert, R.; Weigel, E.; Beyer, D.; Berbalk, M.; Tresp, V. A knowledge graph perspective on supply chain resilience. *arXiv preprint arXiv:2305.08506* **2023**.

Ameri, F.; Wallace, E.; Yoder, R.; Riddick, F. Agri-food supply chain traceability supported by a formal ontology: A grain elevator to processor use case. In Proceedings of the International Design Engineering Technical Conferences and Computers and Information in Engineering Conference. American Society of Mechanical Engineers, 2023, Vol. 87295, p. V002T02A051.

Wang, L.; Liu, X.; Dong, Y.; Zhao, D.; Wang, Z.; Chen, X. Risk evolution analysis of cross-regional water diversion projects based on spatio-temporal knowledge graphs. *Journal of Hydrology* **2025**, *650*, 132533.

Yang, L.; Li, Y.; Tan, J.; Mao, L. Research on risk decision-making generation method for water conservancy project based on multimodal knowledge graph and large language model. *PLoS One* **2025**, *20*, e0330258.

 doi:10.20944/preprints202602.1970.v2

Tang, H.; Feng, J.; Zhou, S. Reservoir Optimization Scheduling Driven by Knowledge Graphs. *Electronics* **2024**, *13*, 2283.

Wang, L.; Liu, X.; Liu, Y.; Li, H.; Liu, J.; Yang, L. Multimodal knowledge graph construction for risk identification in water diversion projects. *Journal of Hydrology* **2024**, *635*, 131155.

Sun, S.; Ding, Y.; Dong, G.; Wang, A. Application of knowledge graph in smart irrigation district management decision making. *Heliyon* **2024**, *10*.

Ospan, A.; Mansurova, M.; Barakhnin, V.; Nugumanova, A.; Titkov, R. The development of a water resource monitoring ontology as a research tool for sustainable regional development. *Data* **2023**, *8*, 162.

He, L.; Ye, W.; Wang, Y.; Feng, H.; Chen, B.; Liang, D. Using knowledge graph and RippleNet algorithms to fulfill smart recommendation of water use policies during shale resources development. *Journal of Hydrology* **2023**, *617*, 128970.

Yan, J.; Gao, Q.; Yu, Y.; Chen, L.; Xu, Z.; Chen, J. Combining knowledge graph with deep adversarial network for water quality prediction. *Environmental Science and Pollution Research* **2023**, *30*, 10360–10376.

Wang, L.H.; Liu, X.M.; Liu, Y.; Li, H.R.; Liu, J.Q.; Yang, L.B. Emergency entity relationship extraction for water diversion project based on pre-trained model and multi-featured graph convolutional network. *Plos one* **2023**, *18*, e0292004.

Wang, L.; Liu, X.; Liu, Y.; Li, H.; Liu, J. Knowledge-driven intelligent recommendation method for emergency plans in water diversion projects. *Journal of hydroinformatics* **2023**, *25*, 2522–2540.

Mezni, H.; Driss, M.; Boulila, W.; Atitallah, S.B.; Sellami, M.; Alharbi, N. Smartwater: A service-oriented and sensor cloud-based framework for smart monitoring of water environments. *Remote Sensing* **2022**, *14*, 922.

Liang, J.; Xie, J.; Wang, X.; Wang, S.; Yu, M. Visualization of multi scenario water resources regulation based on a dualistic water cycle framework. *Water* **2022**, *14*, 1128.

Wang, L.; Hu, Y.; Zheng, Z.; Wu, G.; Lin, J.; Li, J.; Zhang, K. Fine-Grained Dismantling Decision-Making for Distribution Transformers Based on Knowledge Graph Subgraph Contrast and Multimodal Fusion Perception. *Electronics* **2025**, *14*, 2754.

Dang, M.Y.; Wang, Q.C.; Qi, J.; Liu, G.; Li, N.; Chen, W.Q. Green design evaluation of electrical and electronic equipment based on knowledge graph. *ACS Sustainable Chemistry & Engineering* **2023**, *11*, 18011–18020.

Wang, J.; Huang, J.; Li, R. Knowledge graph construction of end-of-life electric vehicle batteries for robotic disassembly. *Applied Sciences* **2023**, *13*, 13153.

Thapa, R.B.; Hernández, D.; Brandt, N.; Klein, J.F.; Hoffmann, E.; Staab, S.; Selzer, M.; Lanza, G. A Roadmap to Create a Knowledge Graph for the Circular Factory for the Perpetual Product **2025**.

Aprilia, A.; Djatna, T.; Indrasti, N.S.; Sugiarto. Deep walk and PCA based conceptual model of sustainable packaging design. In Proceedings of the AIP Conference Proceedings. AIP Publishing LLC, 2023, Vol. 2485, p. 020019.

Yang, Y.; Liu, X.; Tu, X.; Lu, Y.; Wang, Y. Automating the Construction of Environmental Policy Knowledge Graph with Large Language Models. *Sustainability (2071-1050)* **2025**, *17*.

Thimm, H.; Schneider, P. Relation Extraction from Environmental Law Text Using Natural Language Understanding. In Proceedings of the EnviroInfo 2022. Gesellschaft für Informatik eV, 2022, p. 43.

Zhang, Z.; Wang, X.; Meng, L.; Wang, Q. Enterprise development forecast analysis under environmental policy based on knowledge graph. In Proceedings of the IOP Conference Series: Earth and Environmental Science. IOP Publishing, 2021, Vol. 675, p. 012111.

Wang, X.; Meng, L.; Wang, X.; Wang, Q. The construction of environmental-policy-enterprise knowledge graph based on PTA model and PSA model. *Resources, Conservation & Recycling Advances* **2021**, *12*, 200057.

Zheng, C.; Saif, W.; Tang, Y.; Su, X.; Kassem, M. INTELLIGENT KNOWLEDGE GRAPH QUESTION ANSWERING METHOD FOR HEALTH AND SAFETY HAZARD MANAGEMENT USING LARGE LANGUAGE MODELS **2025**.

Cheng, Q.; Zhang, S.; Yang, L. A text feature extraction model for hazardous chemical recovery identification and attribute classification embedded in domain knowledge graph. *Environmental Monitoring and Assessment* **2025**, *197*, 415.

Katzenstein, M.; Etcheverry, L. FAIR and Quality-Aware Air Quality Data Management: A Knowledge Graph-Based Approach **2025**.

Du, W.; Wang, X.; Zhu, Q.; Jing, X.; Liu, X. CPBA-CLIM: An entity-relation extraction model for ontology-based knowledge graph construction in hazardous chemical incident management. *Science progress* **2024**, *107*, 00368504241235510.

 doi:10.20944/preprints202602.1970.v2

Han, F.; Deng, Y.; Liu, Q.; Zhou, Y.; Wang, J.; Huang, Y.; Zhang, Q.; Bian, J. Construction and application of the knowledge graph method in management of soil pollution in contaminated sites: A case study in South China. *Journal of Environmental Management* **2022**, *319*, 115685.

Stagnol, L.; Cherief, A.; Farah, Z.; Le Guenedal, T.; Sakout, S.; Sekine, T. Answering clean tech questions with large language models. *Available at SSRN* **2023**.

Li, H.; Yang, R.; Xu, S.; Xiao, Y.; Zhao, H. Intelligent checking method for construction schemes via fusion of knowledge graph and large language models. *Buildings* **2024**, *14*, 2502.

Wang, M.; Lilis, G.N.; Mavrokapnidis, D.; Katsigarakis, K.; Korolija, I.; Rovas, D. A knowledge graph-based framework to automate the generation of building energy models using geometric relation checking and HVAC topology establishment. *Energy and Buildings* **2024**, *325*, 115035.

Wu, Z.; Wang, Z.; Cheng, J.C.; Kwok, H.H. A knowledge-informed optimization framework for performance-based generative design of sustainable buildings. *Applied Energy* **2024**, *367*, 123318.

Zhang, Z.; Dai, H.; Jiang, D.; Yu, Y. Multi-objective operation rule optimization of wind-solar-hydro hybrid power system based on knowledge graph structure. *Journal of Cleaner Production* **2025**, *486*, 144514.

Hu, Z.; Li, X.; Pan, X.; Wen, S.; Bao, J. A question answering system for assembly process of wind turbines based on multi-modal knowledge graph and large language model. *Journal of engineering design* **2025**, *36*, 1093–1117.

Pang, A.; Wang, S.; Wu, M.; Zhang, J. Automated Report Generation and Knowledge Management System for Photovoltaic Power Stations using Knowledge Graphs. In Proceedings of the 2024 3rd International Conference on Engineering Management and Information Science (EMIS 2024). Atlantis Press, 2024, pp. 360–375.

Popadić, D.; Iglesias, E.; Sakor, A.; Janev, V.; Vidal, M.E. Toward a solution for an energy knowledge graph. In *Semantic Intelligence: Select Proceedings of ISIC 2022*; Springer, 2023; pp. 3–12.

Chatterjee, J.; Dethlefs, N. XAI4Wind: A multimodal knowledge graph database for explainable decision support in operations & maintenance of wind turbines. *arXiv preprint arXiv:2012.10489* **2020**.

Chun, S.; Jung, J.; Jin, X.; Seo, S.; Lee, K.H. Designing an integrated knowledge graph for smart energy services. *Journal of Supercomputing* **2020**, *76*.

Chun, S.; Jin, X.; Seo, S.; Lee, K.H.; Shin, Y.; Lee, I. Knowledge graph modeling for semantic integration of energy services. In Proceedings of the 2018 IEEE international conference on big data and smart computing (BigComp). IEEE, 2018, pp. 732–735.

Zhou, X.; Shi, J.; Dong, L.; Zhang, Y.; Pan, J.; Huang, H. Construction of a Multimodal Knowledge Graph for Power Grid Construction Safety Based on Large Language Models. In Proceedings of the 2024 International Conference on New Power System and Power Electronics (NPSPE). IEEE, 2024, pp. 21–28.

Peng, F.L.; Qiao, Y.K.; Yang, C. Building a knowledge graph for operational hazard management of utility tunnels. *Expert Systems with Applications* **2023**, *223*, 119901.

Wang, X.; El-Gohary, N. Deep learning-based relation extraction and knowledge graph-based representation of construction safety requirements. *Automation in Construction* **2023**, *147*, 104696.

Simone, F.; Ansaldi, S.M.; Agnello, P.; Di Gravio, G.; Patriarca, R. Knowledge in graphs: investigating the completeness of industrial near miss reports. *Safety science* **2023**, *168*, 106305.

Pandithawatta, S.; Ahn, S.; Rameezdeen, R.; Chow, C.W.; Gorjian, N.; Kim, T.W. Development of a knowledge graph for automatic job hazard analysis: The schema. *Sensors* **2023**, *23*, 3893.

Pedro, A.; Pham-Hang, A.T.; Nguyen, P.T.; Pham, H.C. Data-driven construction safety information sharing system based on linked data, ontologies, and knowledge graph technologies. *International journal of environmental research and public health* **2022**, *19*, 794.

Mao, S.; Zhao, Y.; Chen, J.; Wang, B.; Tang, Y. Development of process safety knowledge graph: A Case study on delayed coking process. *Computers & chemical engineering* **2020**, *143*, 107094.

Fang, W.; Ma, L.; Love, P.E.; Luo, H.; Ding, L.; Zhou, A. Knowledge graph for identifying hazards on construction sites: Integrating computer vision with ontology. *Automation in Construction* **2020**, *119*, 103310.

Yang, Y.; Zhang, C.; Song, X.; Dong, Z.; Zhu, H.; Li, W. Contextualized knowledge graph embedding for explainable talent training course recommendation. *ACM Transactions on Information Systems* **2023**, *42*, 1–27.

Loh, C.H.; Wei, S.; Phua, C.T.; Mohd, A.; Tan, A.; Seow, B.K. Nacelle: Knowledge graph-based conversational ai for skills gap analysis to achieve sustainable learning at workplace. In Proceedings of the Proceedings of 2nd International Conference on Artificial Intelligence: Advances and Applications: ICAIAA 2021. Springer, 2022, pp. 359–371.

de Groot, M.; Schutte, J.; Graus, D. Job posting-enriched knowledge graph for skills-based matching. *arXiv preprint arXiv:2109.02554* **2021**.

 doi:10.20944/preprints202602.1970.v2

Yang, J. Knowledge Graph-Based Intelligent Recommendation and Dynamic Matching System for Tobacco Industry Human Resources. In Proceedings of the Proceedings of the 2025 International Conference on Economic Management and Big Data Application, 2025, pp. 478–482.

Al Akasheh, M.; Hujran, O.; Malik, E.F.; Zaki, N. Enhancing the prediction of employee turnover with knowledge graphs and explainable AI. *IEEE Access* **2024**, *12*, 77041–77053.

Nagy, L.; Abonyi, J.; Ruppert, T. Knowledge Graph-Based Framework to Support Human-Centered Collaborative Manufacturing in Industry 5.0. *Applied Sciences* **2024**, *14*, 3398.

Huang, F.; Deng, Y.; Zhang, C.; Guo, M.; Zhan, K.; Sun, S.; Jiang, J.; Sun, Z.; Wu, X. KOSA: KO enhanced salary analytics based on knowledge graph and LLM capabilities. In Proceedings of the 2023 IEEE International Conference on Data Mining Workshops (ICDMW). IEEE, 2023, pp. 499–505.

Wang, Y.; Henni, A.; Fotouhi, A.; Ze, L.M. Leveraging K-hop based Graph for the Staffing Recommender System with Parametric Geolocation. In Proceedings of the 2022 International Conference on Computational Science and Computational Intelligence (CSCI). IEEE, 2022, pp. 664–669.

Aravind Krishnan, A.; Deepak, G. OGDES: An Automatic Ontology Generation Mechanism for Diversity, Equity and Inclusion Studies as a Prospective Domain of Choice Integrating Semantic Intelligence. In Proceedings of the International Conference on Hybrid Intelligent Systems. Springer, 2023, pp. 12–21.

Zheng, G.; Brintrup, A. Enhancing supply chain visibility with generative AI: an exploratory case study on relationship prediction in knowledge graphs. *International Journal of Production Research* **2025**, pp. 1–23.

Verona, L.D.; Lopes, G.R.; Campos, M.L.M. Using Government Data to Uncover Political Power and Influence of Contemporary Slavery Agents in Brazil. In Proceedings of the Workshop on Big Social Data and Urban Computing. Springer, 2018, pp. 123–138.

Zheng, C.; Chen, G.; Chen, H.; Xu, Q.; Zhao, Y.; Zhao, Y.; Yang, Y. Chemical process safety domain knowledge graph-enhanced LLM for efficient emergency response decision support. *The Canadian Journal of Chemical Engineering* **2025**.

Zhao, K.; Lu, X.; Wan, L.; Zhang, L.; Jin, Y.; Wen, P.; Gao, J.; He, M.; Wang, Q.; Zhan, L. From data to insight: Building a knowledge graph for risk analysis of hazardous chemical accidents. *Chinese Journal of Chemical Engineering* **2025**.

Xu, Y.; Chen, G.; Wang, J. KGC-CLIP: Structure-Aware Vision-Language Pre-training for Hazardous Chemical Identification via Knowledge Graph Enhancement. *Available at SSRN 5938365*.

Da Silveira, M.; Deladiennee, L.; Acem, K.; Freudenthal, O. Combining knowledge graphs and LLMs for hazardous chemical information management and reuse. In Proceedings of the 2024 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). IEEE, 2024, pp. 6766–6773.

Myklebust, E.B.; Jiménez-Ruiz, E.; Chen, J.; Wolf, R.; Tollefsen, K.E. Prediction of adverse biological effects of chemicals using knowledge graph embeddings. *Semantic Web* **2022**, *13*, 299–338.

Shin, E.; Yoo, S.; Ju, Y.; Shin, D. Knowledge graph embedding and reasoning for real-time analytics support of chemical diagnosis from exposure symptoms. *Process Safety and Environmental Protection* **2022**, *157*, 92–105.

Zheng, X.; Wang, B.; Zhao, Y.; Mao, S.; Tang, Y. A knowledge graph method for hazardous chemical management: Ontology design and entity identification. *Neurocomputing* **2021**, *430*, 104–111.

Sharma, A.N.; Akbar, K.A.; Thuraisingham, B.; Khan, L. Enhancing Security Insights with KnowGen-RAG: Combining Knowledge Graphs, LLMs, and Multimodal Interpretability. In Proceedings of the Proceedings of the 2025 ACM International Workshop on Security and Privacy Analytics, 2025, pp. 2–12.

Echenim, K.U.; Joshi, K.P. Automating IoT Data Privacy Compliance by Integrating Knowledge Graphs With Large Language Models. *IEEE Access* **2025**.

Qiao, W.; Geng, Z.; Liu, H. KG-PEM: A Data Privacy Protection Assessment Framework Based on Knowledge Graph. In Proceedings of the 2025 10th International Conference on Computer and Communication System (ICCCS). IEEE, 2025, pp. 221–227.

Gilliard, E.; Liu, J.; Aliyu, A.A. Knowledge graph reasoning for cyber attack detection. *IET Communications* **2024**, *18*, 297–308.

Li, P.; Bai, X.; Li, J.; Dong, Y.; Yang, J. Automatic Construction of Knowledge Graph for Personal Sensitive Data. In Proceedings of the International Conference on Frontiers in Cyber Security. Springer, 2023, pp. 252–264.

Liu, J.; Zhan, J. Constructing knowledge graph from cyber threat intelligence using large language model. In Proceedings of the 2023 IEEE International Conference on Big Data (BigData). IEEE, 2023, pp. 516–521.

Echenim, K.U.; Joshi, K.P. Iot-reg: A comprehensive knowledge graph for real-time iot data privacy compliance. In Proceedings of the 2023 IEEE International conference on big data (BigData). IEEE, 2023, pp. 2897–2906.

 doi:10.20944/preprints202602.1970.v2

Gambarelli, G.; Gangemi, A. PRIVAFRAME: a frame-based knowledge graph for sensitive personal data. *Big Data and Cognitive Computing* **2022**, *6*, 90.

Li, Z.; Zeng, J.; Chen, Y.; Liang, Z. AttacKG: Constructing technique knowledge graph from cyber threat intelligence reports. In Proceedings of the European Symposium on Research in Computer Security. Springer, 2022, pp. 589–609.

Sarhan, I.; Spruit, M. Open-cykg: An open cyber threat intelligence knowledge graph. *Knowledge-based systems* **2021**, *233*, 107524.

Kiesling, E.; Ekelhart, A.; Kurniawan, K.; Ekaputra, F. The SEPSES knowledge graph: an integrated resource for cybersecurity. In Proceedings of the International Semantic Web Conference. Springer, 2019, pp. 198–214.

Jia, Y.; Qi, Y.; Shang, H.; Jiang, R.; Li, A. A practical approach to constructing a knowledge graph for cybersecurity. *Engineering* **2018**, *4*, 53–60.

Huang, Y.; Song, X. A Method for Constructing Equipment Fault Diagnosis Knowledge Graph Based on Large Language Model. In Proceedings of the China Simulation Conference. Springer, 2025, pp. 203–217.

Wu, P.; Mou, X.; Gong, L.; Tu, H.; Qiu, L.; Yang, B. An automatic machine fault identification method using the knowledge graph–embedded large language model. *The International Journal of Advanced Manufacturing Technology* **2025**, pp. 1–15.

Wen, P.; Zhang, T.; Hu, Y.; Tao, L. A novel knowledge graph-based identification method for key quality characteristics of complex products. *Expert Systems with Applications* **2025**, p. 129112.

Zhou, B.; Li, X.; Liu, T.; Xu, K.; Liu, W.; Bao, J. CausalKGPT: Industrial structure causal knowledge-enhanced large language model for cause analysis of quality problems in aerospace product manufacturing. *Advanced Engineering Informatics* **2024**, *59*, 102333.

Yang, Y.; Huang, M.; Zhang, C.; Yu, M. Put Theory into Practice Knowledge Graph Based Aviation Quality Reliability Knowledge System. In Proceedings of the 2022 4th International Conference on System Reliability and Safety Engineering (SRSE). IEEE, 2022, pp. 30–34.

Wang, M.; Dong, F.; Liu, H.; Lin, Q. Construction and Application of Knowledge Graph for Aerospace Product Quality Problems. In Proceedings of the 2021 2nd International Conference on Information Science and Education (ICISE-IE). IEEE, 2021, pp. 546–552.

Liu, H.; Dong, F.; Wang, M.; Lin, Q. Modelling and implementation of a knowledge question-answering system for product quality problem based on knowledge graph. In Proceedings of the Journal of Physics: Conference Series. IOP Publishing, 2021, Vol. 2068, p. 012051.

Zhang, J.; Liu, J.; Zhu, Y.; He, F.; Feng, S.; Li, J. Whole-chain supervision method of industrial product quality and safety based on knowledge graph. In Proceedings of the 2021 IEEE International Conference on Industrial Application of Artificial Intelligence (IAAI). IEEE, 2021, pp. 74–78.

Sun, T.; Wang, Q. Multi-source fault detection and diagnosis based on multi-level Knowledge Graph and Bayesian theory reasoning (S). In Proceedings of the SEKE, 2019, pp. 177–232.

He, L.; Jiang, P. Manufacturing knowledge graph: a connectivism to answer production problems query with knowledge reuse. *Ieee Access* **2019**, *7*, 101231–101244.

Li, X.; Wang, X.; Li, G. Aging-Friendly Design Research: Knowledge Graph Construction for Elderly Advantage Applications. *Applied Sciences* **2025**, *15*, 2848.

Li, A.; Han, C.; Xing, X.; Wei, Q.; Chi, Y.; Pu, F. KGSCS—a smart care system for elderly with geriatric chronic diseases: a knowledge graph approach. *BMC Medical Informatics and Decision Making* **2024**, *24*, 73.

Duan, Y.; Ji, P.; Jin, L.; Zou, A.; Yang, J.; Xie, H.; An, N. A knowledge graph for eldercare: Constructing a domain entity graph with guidelines. In Proceedings of the International Conference on Human Aspects of IT for the Aged Population. Springer, 2018, pp. 25–35.

Li, Y.; Zakhozhyi, V.; Zhu, D.; Salazar, L.J. Domain specific knowledge graphs as a service to the public: Powering social-impact funding in the us. In Proceedings of the Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, 2020, pp. 2793–2801.

Gordon, S.N.; Murphy, P.J.; Gallo, J.A.; Huber, P.; Hollander, A.; Edwards, A.; Jankowski, P. People, projects, organizations, and products: Designing a knowledge graph to support multi-stakeholder environmental planning and design. *ISPRS International Journal of Geo-Information* **2021**, *10*, 823.

dAmato, C.; Rubini, G.; Didio, F.; Francioso, D.; Amara, F.Z.; Fanizzi, N. Automated Creation of the Legal Knowledge Graph Addressing Legislation on Violence Against Women: Resource, Methodology and Lessons Learned. *arXiv preprint arXiv:2508.06368* **2025**.

 doi:10.20944/preprints202602.1970.v2

Blin, I.; Stork, L.; Spillner, L.; Santagiustina, C. OKG: A Knowledge Graph for Fine-grained Understanding of Social Media Discourse on Inequality. In Proceedings of the Proceedings of the 12th Knowledge Capture Conference 2023, 2023, pp. 166–174.

Mitra, R.; Dongre, A.; Dangare, P.; Goswami, A.; Tiwari, M.K. Knowledge graph driven credit risk assessment for micro, small and medium-sized enterprises. *International Journal of Production Research* **2024**, *62*, 4273–4289.

Alam, M.N.; Ali, M.M. Loan default risk prediction using knowledge graph. In Proceedings of the 2022 14th International Conference on Knowledge and Smart Technology (KST). IEEE, 2022, pp. 34–39.

Beydon, G.; Suryanto, H.; Guan, C.; Guan, A.; Sugumaran, V. Knowledge Graphs in Support of Credit Risk Assessment **2020**.

Uddin, M.S.; Ahmed, A.; Aktarujjaman, M.; Moniruzzaman, M.; Ahmed, M.; Mridha, M.; Hossen, M.J. A hybrid reinforcement learning and knowledge graph framework for financial risk optimization in healthcare systems. *Scientific Reports* **2025**, *15*, 29057.

Qi, M.; Santos, H.; Pinheiro, P.; McGuinness, D.L.; Bennett, K.P. Demographic and socioeconomic determinants of access to care: A subgroup disparity analysis using new equity-focused measurements. *PLoS One* **2023**, *18*, e0290692.

Talukder, A.K.; Schriml, L.; Ghosh, A.; Biswas, R.; Chakrabarti, P.; Haas, R.E. Diseasomics: Actionable machine interpretable disease knowledge at the point-of-care. *PLOS Digital Health* **2022**, *1*, e0000128.

Huang, L.; Yu, C.; Chi, Y.; Qi, X.; Xu, H. Towards smart healthcare management based on knowledge graph technology. In Proceedings of the Proceedings of the 2019 8th International Conference on Software and Computer Applications, 2019, pp. 330–337.

Jain, A.; Kerboeuf, S.; Barmpounakis, S.; Wendt, S.; Bui, D.T.; Alemany, P.; Nicolicchia, R.; Valero, J.M.J.; Korpi, D.; Moghaddam, M.H.; et al. Knowledge Graph-Based approach for Sustainable 6G End-to-End System Design. *arXiv preprint arXiv:2507.08717* **2025**.

Xiao, H.; Sun, Y. " policies look for the elderly": A knowledge graph based care information recommendation system. In Proceedings of the 2023 26th International conference on computer supported cooperative work in design (CSCWD). IEEE, 2023, pp. 1754–1759.

Ma, W.; Li, M.; Dai, J.; Ding, J.; Chu, Z.; Chen, H. Nutrition-related knowledge graph neural network for food recommendation. *Foods* **2024**, *13*, 2144.

Dang, L.D.; Phan, U.T.; Nguyen, N.T. GENA: A knowledge graph for nutrition and mental health. *Journal of Biomedical Informatics* **2023**, *145*, 104460.

Fu, C.; Pan, X.; Wu, J.; Cai, J.; Huang, Z.; van Harmelen, F.; Zhao, W.; Jiang, X.; He, T. KG4NH: a comprehensive knowledge graph for question answering in dietary nutrition and human health. *IEEE journal of biomedical and health informatics* **2023**.

Fu, C.; Yao, Y.; Wu, J.; Zhao, W.; He, T.; Jiang, X. Multimodal reasoning for nutrition and human health via knowledge graph embedding. In Proceedings of the 2023 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). IEEE, 2023, pp. 1901–1904.

Chen, Y.; Subburathinam, A.; Chen, C.H.; Zaki, M.J. Personalized food recommendation as constrained question answering over a large-scale food knowledge graph. In Proceedings of the Proceedings of the 14th ACM international conference on web search and data mining, 2021, pp. 544–552.

Chi, Y.; Yu, C.; Qi, X.; Xu, H. Knowledge management in healthcare sustainability: a smart healthy diet assistant in traditional Chinese medicine culture. *Sustainability* **2018**, *10*, 4197.

Xu, Q.; Hong, L.; Shen, M.; Yi, B. Disclosing Actual Controller based on Equity Knowledge Graph Learning. In Proceedings of the Proceedings of the 31st ACM SIGKDD Conference on Knowledge Discovery and Data Mining V. 1, 2025, pp. 2735–2744.

Shaw, C.; de Andrade Pereira, F.; de Riet, M.; Hoare, C.; Farghaly, K.; O'Donnell, J. Knowledge graph for policy-and practice-aligned life cycle analysis and reporting. *Automation in Construction* **2025**, *176*, 106282.

Magnanimi, D.; Bellomarini, L.; Ceri, S.; Martinenghi, D. Reactive company control in company knowledge graphs. In Proceedings of the 2023 IEEE 39th International Conference on Data Engineering (ICDE). IEEE, 2023, pp. 3336–3348.

Badmus, O.; Ikumapayi, O.J.; Toromade, R.O.; Sunday, A. Integrating AI-powered knowledge graphs and NLP for intelligent interpretation, summarization, and cross-border financial reporting harmonization. *World Journal of Advanced Research and Reviews* **2025**, *27*, 42–62.

Li, J.; Chang, Y.; Wang, Y.; Zhu, X. Tracking down financial statement fraud by analyzing the supplier-customer relationship network. *Computers & Industrial Engineering* **2023**, *178*, 109118.

 doi:10.20944/preprints202602.1970.v2

Wen, S.; Li, J.; Zhu, X.; Liu, M.; et al. Analysis of financial fraud based on manager knowledge graph. *Procedia computer science* **2022**, *199*, 773–779.

Zehra, S.; Mohsin, S.F.M.; Wasi, S.; Jami, S.I.; Siddiqui, M.S.; Syed, M.K.U.R.R. Financial knowledge graph based financial report query system. *IEEE Access* **2021**, *9*, 69766–69782.

Liu, Y.; Guo, W.; Zhang, H.; Bian, H.; He, Y.; Zhang, X.; Gong, Y.; Dong, J.; Liu, Z. Construction of knowledge graph based on discipline inspection and supervision. In Proceedings of the 2021 IEEE 20th International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom). IEEE, 2021, pp. 1467–1472.

Gao, Y.; Liu, J. Latent Clues Investigation in Discipline Inspection and Supervision Based on Knowledge Graph Analysis. In Proceedings of the 2021 IEEE 24th International Conference on Computer Supported Cooperative Work in Design (CSCWD). IEEE, 2021, pp. 204–209.

Zheng, J.; Li, Y. Machine learning model of tax arrears prediction based on knowledge graph. *Electronic Research Archive* **2023**, *31*.

Lüdemann, N.; Shiba, A.; Thymianis, N.; Heist, N.; Ludwig, C.; Paulheim, H. A knowledge graph for assessing agressive tax planning strategies. In Proceedings of the International Semantic Web Conference. Springer, 2020, pp. 395–410.

Ma, Y.; Shu, Y.; Hu, G. GovGraph: Unveiling Social Governance Innovation Networks with an Integrated Model Framework and Diverse Data for Link Prediction. *Expert Systems with Applications* **2025**, p. 130214.

Tang, Z.; Chi, Y. Research on the Impact of Corporate Serial Emergencies Based on Knowledge Graph. *Procedia Computer Science* **2025**, *266*, 987–995.

Wang, P.; Hu, Q.; Mei, Q.; Wang, S.; Yang, Y.; Guo, D.; Liu, X.; Hu, W.; Chen, J. Intelligent port logistics: A spatiotemporal knowledge graph and AI-agent framework for berth allocation. *Advanced Engineering Informatics* **2025**, *68*, 103633.

Driller, J.; Trang, S.T.N. Unlocking sustainable reporting: Leveraging knowledge graphs for ESG metrics extraction: The role of knowledge graphs in sustainability reporting. In Proceedings of the INFORMATIK 2024. Gesellschaft f¨ur Informatik eV, 2024, pp. 1877–1883.

Raikar, G.V.; Deepak, G. SGKPS: A Semantic AI-Driven Strategy for Knowledge Graph Population for Sustainable Journalism as a domain of choice **2024**.

Benjira, W.; Atigui, F.; Bucher, B.; Grim-Yefsah, M.; Travers, N. Automated mapping between SDG indicators and open data: An LLM-augmented knowledge graph approach. *Data & Knowledge Engineering* **2025**, *156*, 102405.

Wang, Q.; Yu, Z.; Wang, S.; Zhu, Y.; Dai, X.; Zou, Z.; Huang, W.; Claramunt, C. Unveiling optimal SDG pathways: an innovative automated recommendation approach integrating graph pruning, intent graph, and attention mechanism. *International Journal of Digital Earth* **2025**, *18*, 2513048.

Lin, Y.D.; Liao, G.Z. Large Language Model-Based Knowledge Graph System Construction for Sustainable Development Goals: An AI-Based Speculative Design Perspective. *arXiv preprint arXiv:2504.12309* **2025**.

Mishra, P.; Narayanasamy, S.K.; Srinivasan, K. Context-Aware Embedded Language Transformers for Evaluating Climate Change based Sustainable Development Goals. *IEEE Access* **2025**.

Kilanioti, I.; Papadopoulos, G.A. AI-based knowledge graph construction and distributed storage for collaboration on the Sustainable Development Goals. In Proceedings of the Proceedings of the 13th Hellenic Conference on Artificial Intelligence, 2024, pp. 1–6.

Benjira, W.; Atigui, F.; Bucher, B.; Grim-Yefsah, M.; Travers, N. Web Open Data to SDG Indicators: Towards an LLM-Augmented Knowledge Graph Solution. In Proceedings of the International Conference on Web Information Systems Engineering. Springer, 2024, pp. 90–100.

DeBellis, M.; Arellano, C.; Guo, P.; Jyothi, T.; Suresh, V.; Kron, K. DaanKG: An Ontology model of the UN Sustainable Development Goals to Facilitate and Improve Corporate Social Responsibility. In Proceedings of the JOWO, 2023.

Mandilara, I.; Fotopoulou, E.; Mariarona, C.; Zafeiropoulos, A.; Papavassiliou, S. Knowledge Graph Data Enrichment based on a Software Library for Text Mapping to the Sustainable Development Goals. In Proceedings of the TEXT2KG/BiKE@ ESWC, 2023, pp. 51–69.

Orellana, D.F.P.; Piedra, N. Semantic Enrichment of Open Dataset related to sustainable Development Goals using Open Knowledge Graphs. In Proceedings of the 2021 XVI Latin American Conference on Learning Technologies (LACLO). IEEE, 2021, pp. 470–473.

Eguiguren, J.; Piedra, N. Description of Open Data Sets as Semantic Knowledge Graphs to Contribute to Actions Related to the 2030 Agenda and the Sustainable Development Goals. In Proceedings of the Iberoamerican Knowledge Graphs and Semantic Web Conference. Springer, 2020, pp. 181–194.

 doi:10.20944/preprints202602.1970.v2

Eguiguren, J.E.; Piedra, N. Connecting Open Data and Sustainable Development Goals using a Semantic Knowledge Graph approach **2019**.

Zhou, Y.; Cao, Y.; Perzylo, A. Towards digital sustainability reporting: An ontology for mapping of indicators in gri and esrs. In *Knowledge Graphs in the Age of Language Models and Neuro-Symbolic AI*; IOS Press, 2024; pp. 191–207.

Gupta, T.K.; Goel, T.; Verma, I.; Dey, L.; Bhardwaj, S. Knowledge graph aided llm based esg question-answering from news. In Proceedings of the Proceedings of the 2nd International Workshop on Knowledge Graphs for Sustainability (KG4S2024), co-located with ESWC, 2024, Vol. 3753.

Koloski, B.; Montariol, S.; Purver, M.; Pollak, S. Knowledge informed sustainability detection from short financial texts. In Proceedings of the Proceedings of the fourth workshop on financial technology and natural language processing (FinNLP), 2022, pp. 228–234.

Belova, A.; Gallina, V.; Revenko, A.; Ahmeti, A. Digital Product Passport: Initial System Architecture with Knowledge Graphs and Data Spaces **2025**.

Marconnet, B.; Gaha, R.; Eynard, B. Context-Aware Sustainable Design: Knowledge Graph-Based Methodology for Proactive Circular Disassembly of Smart Products. In Proceedings of the Technological Systems, Sustainability and Safety, 2024.

Vasileiadis, M.; Mexis, K.; Trokanas, N.; Dalamagas, T.; Papageorgiou, T.; Kokossis, A. Leveraging Semantics and Machine Learning to Automate Circular Economy Operations for the Scrap Metals Industry. In *Computer Aided Chemical Engineering*; Elsevier, 2023; Vol. 52, pp. 3441–3446.

Kebede, R.; Moscati, A.; Tan, H.; Johansson, P. Circular economy in the built environment: a framework for implementing digital product passports with knowledge graphs. In Proceedings of the EC3 Conference 2023. European Council on Computing in Construction, 2023, Vol. 4, pp. 0–0.

Bauer, D.; Longley, T.; Ma, Y.; Wilson, T. NLP in Human Rights Research–Extracting Knowledge Graphs About Police and Army Units and Their Commanders. *arXiv preprint arXiv:2201.05230* **2022**.

Zhao, Y.; Dai, C.; Zhuo, W.; Fu, T.C.; Xiu, Y.; Niyato, D.; Low, J.Z.; Zhuang, E.H.H.; Tan, D.Z.L. AGENTICT ^{2} S: Robust Text-to-SPARQL via Agentic Collaborative Reasoning over Heterogeneous Knowledge Graphs for the Circular Economy. *arXiv preprint arXiv:2508.01815* **2025**.

Karimanzira, D.; Rauschenbach, T.; Hellmund, T.; Ritzau, L. Improved Flood Management and Risk Communication Through Large Language Models. *Algorithms* **2025**, *18*, 713.

Ong, L.; Karmakar, G.; Atherton, J.; Zhou, X.; Lim, M.Q.; Chadzynski, A.; Li, L.; Wang, X.; Kraft, M. Embedding energy storage systems into a dynamic knowledge graph. *Industrial & Engineering Chemistry Research* **2022**, *61*, 8390–8398.

Janowicz, K.; Hitzler, P.; Li, W.; Rehberger, D.; Schildhauer, M.; Zhu, R.; Shimizu, C.; Fisher, C.; Cai, L.; Mai, G.; et al. Know, Know Where, KnowWhereGraph: A densely connected, cross-domain knowledge graph and geo-enrichment service stack for applications in environmental intelligence. *AI Magazine* **2022**, *43*, 30–39.

Hisano, R.; Sornette, D.; Mizuno, T. Prediction of ESG compliance using a heterogeneous information network. *Journal of Big Data* **2020**, *7*, 22.

De Mulder, G.; de Vleeschauwer, E.; De Meester, B.; Colpaert, P.; Hartig, O. The Open Circularity Platform: a Decentralized Data Sharing Platform for Circular Value Networks. In Proceedings of the 2nd International Workshop on Knowledge Graphs for Sustainability (KG4S 2024) colocated with the 21st Extended Semantic Web Conference (ESWC 2024), 2024, Vol. 3753, pp. 42–52.

Gautam, N.; Shumway, D.; Kowalcyk, M.; Khanal, S.; Caragea, D.; Caragea, C.; Mcginty, H.; Dorevitch, S. Leveraging existing literature on the web and deep neural models to build a knowledge graph focused on water quality and health risks. In Proceedings of the Proceedings of the ACM Web Conference 2023, 2023, pp. 4161–4171.

He, F.; Fan, J.; Deng, Y.; Zhang, X.; Lau, K.T.; Wang, D. Smart metering data enhancement in sustainable buildings via knowledge graph-guided graph neural networks. *Knowledge-Based Systems* **2025**, p. 114016.

Li, J.; Sun, Y.; Li, C.; Hu, Y.; Wang, C. Industry chain graph building based on text semantic association mining. In Proceedings of the 2021 International Joint Conference on Neural Networks (IJCNN). IEEE, 2021, pp. 1–8.

Szekely, P.; Knoblock, C.A.; Slepicka, J.; Philpot, A.; Singh, A.; Yin, C.; Kapoor, D.; Natarajan, P.; Marcu, D.; Knight, K.; et al. Building and using a knowledge graph to combat human trafficking. In Proceedings of the International Semantic Web Conference. Springer, 2015, pp. 205–221.

Brockmann, N.; Elson Kosasih, E.; Brintrup, A. Supply chain link prediction on uncertain knowledge graph. *ACM SIGKDD Explorations Newsletter* **2022**, *24*, 124–130.

 doi:10.20944/preprints202602.1970.v2

Vijaya, A.; Qadri, F.D.; Angreani, L.S.; Wicaksono, H. ESGOnt: An ontology-based framework for Enhancing Environmental, Social, and Governance (ESG) assessments and aligning with Sustainable Development Goals (SDG). *Resources, Environment and Sustainability* **2025**, p. 100262.

Kilanioti, I.; A. Papadopoulos, G. A knowledge graph-based deep learning framework for efficient content similarity search of Sustainable Development Goals data. *Data Intelligence* **2023**, *5*, 663–684.

Ji, Z.; Wang, X.; Zhang, J.; Wu, D. Construction and application of knowledge graph for grid dispatch fault handling based on pre-trained model. *Global Energy Interconnection* **2023**, *6*, 493–504.

Shi, Y.; Zhou, K.; Zhou, M.; Li, S.; Liu, W. Temporal knowledge graph for food risk prediction. *IEEE Transactions on Artificial Intelligence* **2023**, *5*, 2217–2226.

Yan, J.; Jiao, H.; Zhang, H.; Yu, Y.; Chen, J.; Zhou, S. Clustering Analysis of Water Usage Structures Based on Knowledge Graph. In Proceedings of the 2023 4th International Conference on Computer, Big Data and Artificial Intelligence (ICCBD+ AI). IEEE, 2023, pp. 241–250.

Diamantini, C.; Potena, D.; Storti, E. SemPI: A semantic framework for the collaborative construction and maintenance of a shared dictionary of performance indicators. *Future Generation Computer Systems* **2016**, *54*, 352–365.

Yoo, S.; Shin, E.; Shin, D. SEARCH: A Symptom-based Expert for Advanced Response to Chemical Hazards. In *Computer Aided Chemical Engineering*; Elsevier, 2021; Vol. 50, pp. 1029–1034.