# A Two-Stage Pipeline for Knowledge Graph-Driven Business Strategy

Lucia Siciliani[1], Eleonora Ghizzota[1,*], Alex Jordan[2,3], Alessandro Petruzzelli[1], Giuseppe Spillo[1], Pierpaolo Basile[1], Davide Sola[2,3], Pasquale Lops[1], Giovanni Scarso Borioli[2,3] and Giovanni Semeraro[1]

[1]*Department of Computer Science, University of Bari Aldo Moro, via E. Orabona n.4, Bari, Italy*
[2]*ESCP Business School, London, UK*
[3]*3HORIZONS, London, UK*

**Abstract**

Business strategy requires integrating organization-specific information with broader domain knowledge to support grounded and informed decision-making. Knowledge Graphs provide a suitable representation for organizing such heterogeneous information, but constructing comprehensive business KGs from unstructured sources remains challenging. In this paper, we propose a two-stage pipeline for knowledge graph-driven business strategy that combines document-driven graph generation with ontology-guided graph construction. In the first stage, an initial knowledge graph is automatically generated from documents provided by the organization. This graph captures the entities, relationships, and thematic structures emerging directly from the source material, providing a structured representation of the organization's available knowledge. In the second stage, the document-derived graph serves as contextual evidence to construct a domain-specific business strategy Knowledge Graph. A strategy ontology, developed from expert knowledge, guides the identification, validation, and organization of relevant business entities and relationships, including companies, competitors, offerings, customer segments, markets, and strategic dynamics. This second stage therefore transforms the information contained in the initial graph into an ontology-compliant representation tailored to strategic analysis. By separating general-purpose knowledge extraction from expert-guided domain modelling, the proposed methodology combines the flexibility of data-driven graph generation with the consistency and interpretability of structured domain knowledge. The resulting graph integrates organization-specific evidence and expert-defined strategic concepts, providing a structured knowledge base for downstream retrieval, reasoning, and generative applications. We demonstrate the proposed methodology in the business strategy domain, illustrating how the two-stage process supports the construction of company-specific KGs from heterogeneous organizational documents.

**Keywords**

Large Language Model, Knowledge Graph, Retrieval Augmented Generation, Business Strategy

## 1. Introduction

Large Language Models (LLMs) are increasingly being adopted to support complex knowledge-intensive tasks across multiple domains [1]. While LLMs provide unprecedented capabilities for processing unstructured text, their reasoning remains limited by the lack of explicit, structured representations of domain knowledge and by the difficulty of integrating information originating from multiple sources [2].

Knowledge Graphs (KGs) offer a natural solution to this challenge by organizing information into structured entities and relationships that explicitly capture the semantics of a domain. Besides improving interpretability, KGs provide an effective substrate for graph-based retrieval and reasoning, enabling

*RAGE-KG 2026: Workshop on Retrieval-Augmented Generation Enabled by Knowledge Graphs, October 25–26, 2026, Bari, Italy*
*Corresponding author.
lucia.siciliani@uniba.it (L. Siciliani); e.ghizzota@phd.uniba.it (E. Ghizzota); ajordan@3horizons.com (A. Jordan); alessandro.petruzzelli@uniba.it (A. Petruzzelli); giuseppe.spillo@uniba.it (G. Spillo); pierpaolo.basile@uniba.it (P. Basile); dsola@3horizons.com (D. Sola); pasquale.lops@uniba.it (P. Lops); gsb@3horizons.com (G. S. Borioli); giovanni.semeraro@uniba.it (G. Semeraro)
0000-0002-1438-280X (L. Siciliani); 0000-0002-0751-3891 (E. Ghizzota); 0009-0008-2880-6715 (A. Petruzzelli); 0000-0001-8345-4232 (G. Spillo); 0000-0002-0545-1105 (P. Basile); 0009-0000-3530-3173 (D. Sola); 0000-0002-7185-3014 (P. Lops); 0000-0001-6883-1853 (G. Semeraro)
© 2026 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR Workshop Proceedings ceur-ws.org ISSN 1613-0073
published 2026-11-04

CEUR-WS.org/Vol-4312/paper24.pdf

Retrieval-Augmented Generation (RAG) systems to ground LLM responses on structured knowledge rather than relying solely on parametric memory [3]. As a result, Graph Retrieval-Augmented Generation (GraphRAG) [4] has recently emerged as a promising paradigm for improving the reliability and explainability of LLM-based systems.

Despite these advances, constructing domain-specific knowledge graphs remains a significant challenge [5, 6]. Recent approaches have focused on automatically generating graphs from unstructured documents, producing representations that faithfully capture the information explicitly contained in the source material [3]. While these approaches effectively model organizational knowledge, they are inherently limited to the facts expressed in the available documents and often fail to capture the conceptual abstractions required by specialized domains.

Ontology-driven approaches address this limitation by relying on explicit conceptual models that represent complex domain knowledge. However, they typically require substantial manual effort to design, maintain, and evolve the extraction pipelines needed to populate the graph. This challenge is especially relevant in business strategy, which is characterized by multiple schools of thought and evolving theories [7]. Concepts such as competitive advantage, value proposition, customer segmentation, or market positioning are interpreted differently depending on the strategic framework being adopted. Consequently, it is difficult to assume a universally accepted ontology as a common foundation for strategic reasoning. Any knowledge graph intended to support strategic reasoning must therefore reflect the conceptualization adopted by the business experts themselves.

Consider a company operating in the B2B foodservice desserts market. Its internal documentation describes products, manufacturing processes, customers, suppliers, and commercial activities, enabling the construction of a graph representing the organization's factual knowledge. However, many of the concepts required for strategic analysis, such as competitors, market positioning, strategic capabilities, customer segments, value propositions, and market dynamics, are either only partially represented or entirely absent from the available documents. These concepts cannot simply be extracted because they emerge from expert interpretation rather than from explicit textual evidence. Building a knowledge graph suitable for business strategy therefore requires combining organization-specific information with an explicit strategic conceptualization provided by domain experts. These observations reveal a gap: current methods do not explicitly support the progressive transformation of organization-specific, document-derived knowledge into a business strategy knowledge graph grounded in an expert-defined view of the domain.

In this paper, we propose a two-stage pipeline for knowledge graph-driven business strategy that bridges these two paradigms. Rather than constructing the final business knowledge graph directly from unstructured documents, our methodology progressively transforms organizational knowledge through two complementary stages. The first stage automatically generates an intermediate graph from the organization's documents, producing a structured representation of the factual knowledge explicitly contained in the available data. The second stage uses this intermediate graph as contextual evidence to construct a business strategy knowledge graph, guided by a domain ontology developed by strategy experts. Rather than serving solely as a schema, the ontology explicitly encodes the conceptual assumptions adopted by the experts and guides the identification of strategic entities and relationships, thereby enabling the integration of organization-specific evidence with domain knowledge into a coherent, semantically consistent representation.

The proposed methodology separates the extraction of factual organizational knowledge from its interpretation through a strategic conceptual model. This distinction provides several advantages. First, it preserves the provenance of the information extracted from company documents while allowing the final graph to incorporate concepts that extend beyond the original textual evidence. Second, it enables the business knowledge graph to evolve as the strategic conceptualization changes without requiring the document extraction process to be redesigned. Finally, the resulting graph provides a richer foundation for downstream applications such as GraphRAG, strategic analysis, and decision support.

## 2. Related Work

Business strategy requires integrating heterogeneous sources of information concerning markets, competitors, customers, internal capabilities, and organizational objectives. While established frameworks such as PESTEL, SWOT, and Porter's competitive model provide valuable conceptual foundations for strategic analysis [8], their practical application remains largely dependent on expert knowledge [9]. Recent work has explored the use of LLMs to support strategic reasoning and decision-making [10, 11]. However, LLMs operate primarily on unstructured text, making them prone to hallucinations and limiting their ability to explicitly represent the complex relationships that underlie business strategy [12]. These limitations motivate the use of KGs as structured semantic representations that support transparent reasoning and retrieval.

The recent success of LLMs has also transformed knowledge graph construction. Existing research has investigated LLM-assisted ontology engineering [13, 14, 15], automated ontology generation from textual resources [16, 17], and LLM-based knowledge extraction and graph population [18, 19]. Recent surveys organize this rapidly growing literature around ontology engineering, knowledge extraction, and knowledge fusion, while also distinguishing schema-based approaches, which prioritize consistency and normalization, from schema-free methods focused on open discovery [20]. For example, the framework proposed in [19] first extracts relational triples, then defines and canonicalizes their schema, and can operate with either an existing or a self-generated schema. Despite these advances, most approaches either construct graphs directly from text, induce a schema from extracted content, or use an existing ontology mainly as the target structure for population.

A closely related direction concerns ontology-guided knowledge extraction, where the ontology actively constrains LLM outputs. OntoLogX combines retrieval, structured generation, iterative correction, and ontology validation to construct ontology-compliant knowledge graphs from cybersecurity logs [21]. ODKE+ dynamically generates ontology snippets tailored to entity types and uses them to guide large-scale fact extraction, grounding, normalization, and ingestion [22]. These systems demonstrate the value of ontology-guided prompting and validation, but they focus on domain-specific extraction workflows. In our approach, the ontology acts as the single source of truth from which structured extraction schemas, validation constraints, and graph-population logic are automatically derived. Moreover, the ontology-guided stage operates over an intermediate graph generated from organizational documents rather than directly over isolated textual passages.

Finally, Retrieval-Augmented Generation (RAG) grounds LLM outputs in external knowledge retrieved at inference time [23]. Conventional RAG generally relies on textual chunks and vector similarity, which limits its ability to preserve structural dependencies and retrieve evidence distributed across related concepts. Graph Retrieval-Augmented Generation (GraphRAG) addresses these limitations by leveraging graph-structured representations for both retrieval and generation [3, 4, 24]. Recent surveys distinguish knowledge-based, index-based, and hybrid GraphRAG architectures, reflecting whether graphs act as knowledge carriers, retrieval indices, or both [25]. Existing GraphRAG research, however, largely assumes that a suitable graph is already available and focuses on retrieval and reasoning over it.

Although recent research has significantly advanced ontology engineering, ontology-guided extraction, and GraphRAG independently, comparatively little attention has been devoted to methodologies that integrate these components into a unified graph generation pipeline. Our work addresses this gap by proposing a two-stage methodology that first captures organizational evidence in a document-derived graph and then transforms it into a schema-driven *Business Strategy Knowledge Graph* for downstream strategic question answering.

## 3. Methodology

Our objective is to construct a *Business Strategy Knowledge Graph* that integrates organization-specific information with expert knowledge into a unified semantic representation for strategic analysis and GraphRAG applications. Rather than generating this graph directly from unstructured documents, we

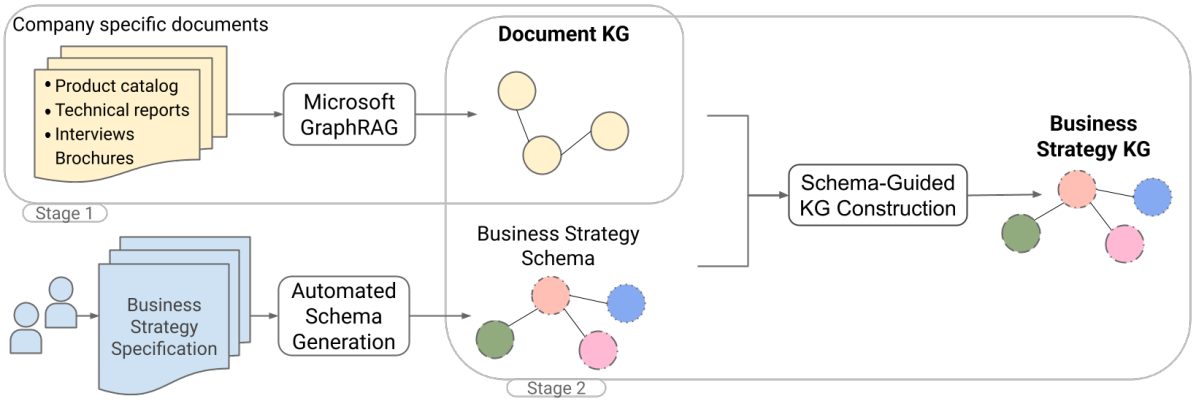



**Figure 1:** Overview of the proposed two-stage pipeline for constructing a Business Strategy Knowledge Graph from company documents and expert knowledge.

adopt a two-stage graph construction methodology.

In the first stage, an initial *Document KG* is generated from the organization's documents, capturing the entities, relationships, and semantic structures explicitly described in the available material. While this graph faithfully represents the documented knowledge, it lacks many concepts required for strategic reasoning, such as competitors, market segments, value propositions, and industry trends. The second stage enriches the *Document KG* into a *Business Strategy KG* using a domain schema developed by business experts. The schema defines the target conceptual structure and guides the extraction and organization of strategic entities and relationships, using the document graph as contextual evidence

In this work, the term *business strategy schema* refers to the machine-readable schema automatically compiled from an expert-authored business strategy specification. The strategic conceptualization itself is not generated automatically: domain experts define the relevant concepts, semantic definitions, properties, and admissible relationships. Instead, automation is applied to the implementation layer, where the specification is translated into structured extraction schemas, validation rules, prompts, and graph-population logic. The implementation artifacts used in this work are available in the companion GitHub repository [1].

Figure 1 illustrates the overall methodology. The *Document KG* generated during the first stage acts as an intermediate semantic representation that bridges raw textual information and the final schema-compliant *Business Strategy KG*. This separation allows factual organizational knowledge and expert-driven strategic knowledge to be integrated while maintaining a clear distinction between evidence extracted from company documents and concepts introduced through the business strategy schema.

Throughout this section, we illustrate the proposed methodology using Maison Clémont, a company operating in the B2B foodservice desserts market, supplying frozen and chilled desserts to restaurants, hotels, caterers, and other foodservice operators. While the company's documentation describes products, production processes, commercial offerings, and customer relationships, strategic analysis also requires understanding its competitors, target markets, value propositions, and positioning within the broader foodservice ecosystem. This example highlights the need to progressively enrich document-derived knowledge into a structured business strategy representation, as well as the practical grounding of the proposed work.

### 3.1. Stage 1: Document-Driven Graph Generation

The first stage of the proposed pipeline aims to transform the heterogeneous documents provided by an organization into a structured graph representation that captures the knowledge explicitly contained in the available material. Rather than reasoning directly over unstructured text throughout the entire

[1] https://github.com/strategyinaction/kgbs_ragekg

pipeline, we first construct an intermediate *Document KG* that organizes the extracted information into interconnected entities and relationships. This KG serves as the factual foundation for the subsequent construction of the *Business Strategy KG*.

**Graph Generation.** For generating the *Document KG* we employed Microsoft GraphRAG[2] [3], a graph-based RAG strategy for *sense-making* tasks over an entire text corpus. Its end-to-end pipeline consists of two main phases, indexing and querying, but we only leverage the indexing functionality. It accepts heterogeneous organizational documents that describe different aspects of the company, including technical documentation, product descriptions, commercial materials, corporate reports, interviews, and other business documents. Since these sources originate from different departments and are typically produced for operational rather than strategic purposes, they contain fragmented and partially overlapping information. Prior to graph construction, the documents are normalized into a common textual representation[3], enabling a uniform processing pipeline.

During the indexing stage, the input documents are transformed into an initial *Document KG* that captures the information explicitly contained in the company documentation. The corpus is first divided into text units, which are processed by a Large Language Model (LLM) to identify entities and semantic relationships. Rather than producing isolated entities or triples, the extracted information is organized into an interconnected graph that preserves the contextual relationships identified across the documents. To support semantic retrieval, text embeddings are also generated for the extracted textual units. In our implementation, we employ OpenAI GPT-5-mini for knowledge extraction and OpenAI text-embedding-3-small for embedding generation. Graph extraction is performed using the default GraphRAG indexing pipeline, which can automatically adapt its extraction prompts to a target domain through prompt auto-tuning. We configure the system using the domain descriptor *"business strategy"*, allowing GraphRAG to generate domain-specific task descriptions and few-shot examples without requiring manually engineered prompts.

The resulting graph is hierarchically clustered using the Leiden algorithm [26]. Community summaries are generated bottom-up, aggregating entities, relationships, and claims from lower-level communities into higher-level reports. Once the indexing stage is completed, the resulting knowledge graph is imported into Neo4j[4] [27], a native graph database queried through Cypher (see Figure 2).

Using the running example introduced before, the available documentation for Maison Clémont includes information describing the company's dessert portfolio, manufacturing capabilities, target customers, and commercial activities. While these documents provide valuable knowledge about the organization, they do not explicitly represent the strategic concepts required for business analysis: the generated *Document KG* is intentionally domain-independent. Its purpose is not to model business strategy but rather to faithfully represent the information explicitly available in the source documentation. Consequently, the graph reflects the terminology and concepts employed within the organization without imposing an external conceptual model: the Maison Clémont *Document KG* contains entities referring to products (e.g., frozen desserts), organizational units, customers (restaurants, hotels, caterers), manufacturing activities, ingredients, certifications, and commercial offerings, together with the relationships explicitly supported by the available documentation.

The graph generation process yields 17,039 nodes, of which 156 are chunks, 5,612 entities, 1,335 communities, 9,926 findings, and 10 documents. Figure 2 shows a portion of the *Document KG*. Yellow nodes are Document nodes, linked to Chunk green nodes via the Part Of edge. Chunk nodes are linked to Community red nodes via Has Chunk relation. Since communities are strictly hierarchical, their hierarchy is made explicit with edges Has Community Parent and Has Community Child between two different Community nodes. Each Chunk node is also linked to its respective blue Entity nodes via the Has Entity relation. These are the entities occurring in the text chunks, identified during the generation process. The generic Entity label is further specialized, e.g., franchise affiliation, dietary requirement, consumer decision driver, location. Entity nodes are related to Community

[2] https://microsoft.github.io/graphrag/
[3] We use Docling (https://www.docling.ai/) for extracting text from documents.
[4] https://neo4j.com/

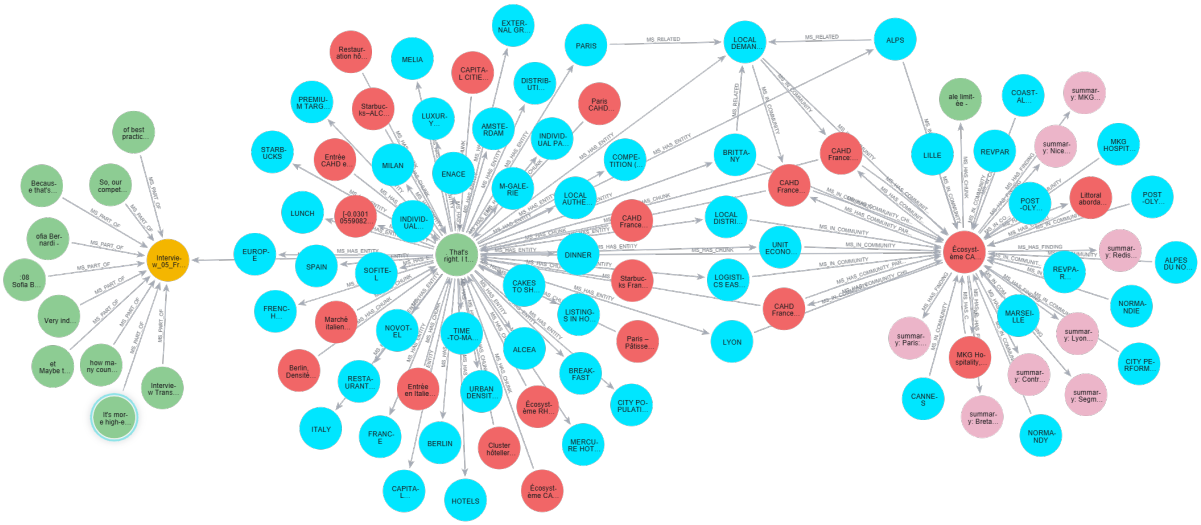


**Figure 2:** A portion of the *Document Knowledge Graph*.

nodes via In Community edge; Entity nodes can be linked to other Entity via Related edge. Finally, Community nodes can have related Finding pink nodes, linked via Has Finding. Findings are report-style summaries for each community in the hierarchical graph using a scalable approach suitable for large datasets. These summaries provide a high-level overview of the dataset's global structure and semantic organization, making them valuable even in the absence of a specific query. For instance, users can browse summaries at higher levels of the hierarchy to identify broad themes and then explore lower-level community reports for more detailed insights into specific subtopics. In this work, however, we primarily leverage these summaries as components of a graph-based index for answering global queries.

The *Document KG* constitutes an intermediate semantic representation within the proposed methodology. Instead of querying the original documents during the second stage, the *Business Strategy KG* is constructed by retrieving contextual evidence from this graph. Using the intermediate graph provides several advantages. First, it consolidates information originating from multiple documents into a unified representation, reducing redundancy and enabling evidence to be retrieved through semantic relationships rather than isolated textual passages. Secondly, it exposes explicit connections between concepts that would otherwise remain distributed across different documents. Furthermore, it provides a structured context that can be exploited during ontology-guided knowledge extraction, allowing the second stage to reason over previously extracted organizational knowledge rather than repeatedly processing the original corpus. This design also establishes a clear separation of responsibilities between the two stages of the pipeline. The first stage focuses exclusively on extracting and organizing factual organizational knowledge, while the second stage is responsible for interpreting this knowledge according to the business strategy schema.

While the *Document KG* effectively captures an organization's explicitly documented facts, such as its products, customers, and operational activities, it is insufficient for strategic reasoning on its own. Because it relies strictly on source texts, it lacks critical business analysis concepts like competitors, market positioning, segmentation, and market dynamics. Additionally, its entities reflect the layout of the source documents rather than a domain-specific model, restricting interoperability and advanced analysis. To address these limitations, the proposed methodology includes a second stage that transforms the intermediate *Document KG* into an ontology-compliant *Business Strategy KG*, successfully merging company evidence with expert strategic knowledge.

### 3.2. Stage 2: Business Strategy Knowledge Graph Construction

The objective of the second stage is therefore to transform this intermediate representation into a *Business Strategy Knowledge Graph* guided by a business strategy schema developed with domain experts. Rather than requiring experts to design formal schemas or ontologies or manually implement extraction pipelines, the proposed framework adopts a lightweight Markdown specification. Experts can define the business concepts relevant for strategic analysis together with their semantic definitions, properties, and relationships in a human-readable document. This specification serves as a source of truth for the entire graph construction process. Thus, the proposed framework does not replace expert ontology design: it is aimed at reducing the engineering effort required to operationalize an expert-defined conceptual model.

Starting from this specification, the framework automatically generates the *Business Strategy Schema*, which defines the entities relevant for strategic analysis (e.g., Company, Competitors, Offering), together with their properties, admissible relationships, and validation constraints. The same generation process also derives the structured extraction schemas, validation rules, and graph population logic required to instantiate the final knowledge graph. This automated workflow decouples domain modeling from implementation. Domain experts only need to maintain the Markdown specification, while any modification to concepts, properties, or relationships is automatically propagated to the extraction pipeline. Consequently, the populated graph remains fully aligned with the experts' conceptual model without requiring manual updates to the extraction logic.

The *Business Strategy Knowledge Graph* is then constructed iteratively by processing the concepts defined in the generated schema. For each entity type, the pipeline retrieves the most relevant contextual evidence from the intermediate *Document KG*, whose interconnected entities and relationships provide richer context than isolated document passages.

**Context Retrieval.** Information relevant to the graph construction query is retrieved from the *Document KG* via a custom Graph RAG Neo4j pipeline[5] that uses a Cypher implementation of Microsoft GraphRAG Local search[6] [3]. Starting from the entities that are semantically most similar to the extraction query, the retrieval process gathers additional contextual evidence, including neighboring entities, relationships, and community reports. To improve efficiency, retrieval is restricted to the top-$k$ entry nodes ($k = 3$) and their respective top-$n$ neighboring nodes ($n = 1$), avoiding exploration of the entire *Document KG* while preserving the most informative contextual evidence.

**Schema-Guided KG Construction.** The retrieved evidence is transformed into natural language and combined with the ontology definition of the target business concept to construct the prompt provided to the LLM. In parallel, the same ontology definition is automatically compiled into a structured extraction schema specifying the expected properties, semantic meaning, and structural constraints of the generated instances. Consequently, the ontology plays a dual role: it guides the LLM (GPT-5-mini though the OpenAI APIs) during extraction through concept-specific prompts while simultaneously constraining the generated output through schema validation. Rather than generating unconstrained business knowledge, the LLM instantiates ontology-compliant business concepts with the retrieved evidence. Each instance is subsequently inserted into the *Business Strategy KG*.

After all entity types have been instantiated, the pipeline constructs the semantic relationships connecting them. For each relationship type, the generated schema specifies the admissible source and target concepts together with the semantic meaning of the relationship. Candidate entities are retrieved from the graph, while supporting evidence is obtained from the *Document KG*. The LLM evaluates the available evidence to determine whether a relationship is justified and generates the corresponding semantic links, which are subsequently validated against the Business Strategy Schema before being materialized in the final KG. By separating entity construction from relationship generation, the pipeline progressively builds a coherent business strategy representation while preserving semantic consistency throughout the extraction process.

The resulting *Business Strategy KG* integrates three complementary sources of knowledge: (i) the

[5] https://neo4j.com/docs/neo4j-graphrag-python/current/user_guide_rag.html
[6] https://microsoft.github.io/graphrag/query/local_search/

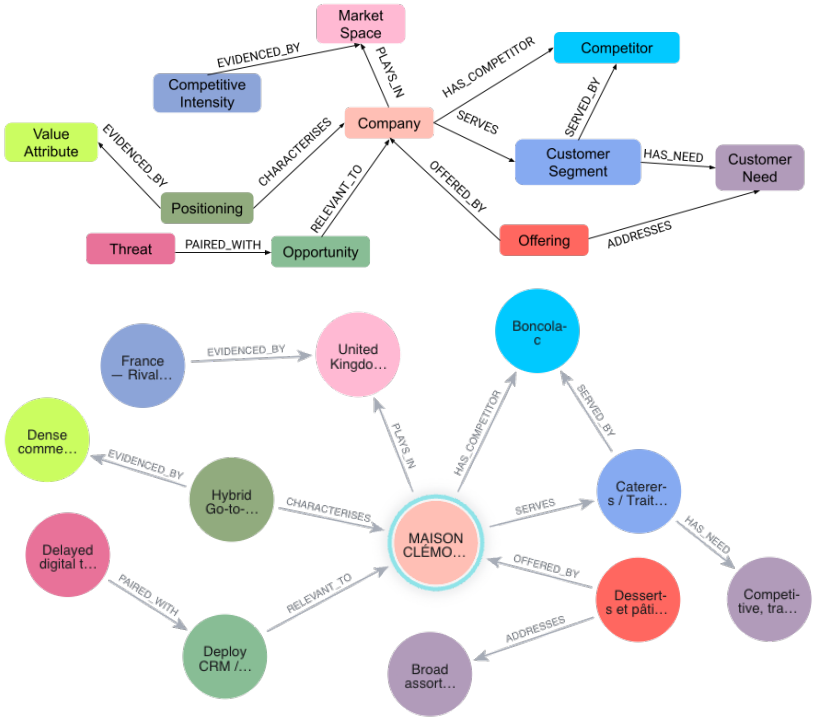



**Figure 3:** Top: portion of the Business Strategy Schema. Bottom: portion of the final Business Strategy KG. Summary statistics of the complete graph are reported in Table 1.

**Table 1**
Final Graph Statistics for Maison Clémont.

| | Document KG | Business Strategy KG |
|---|---|---|
| Instantiated Classes | 24 | 70 |
| Instantiated Relationship Types | 8 | 140 |
| Distinct Node Properties | 18 | 253 |
| Distinct Relationship Properties | 7 | 457 |
| Node Instances | 17,039 | 602 |
| Relationship Instances | 51,964 | 4007 |

factual information extracted from the company's documentation during the *Document KG* generation stage, (ii) the conceptual organization provided by the expert-defined business strategy schema, and (iii) additional contextual knowledge acquired through external retrieval with the LLM (e.g. web search), allowing the graph to be enriched with relevant strategic information not explicitly contained in the company's documentation.

Returning to the Maison Clémont example, the initial *Document KG* contains entities describing products, customers, production activities, and commercial information directly extracted from the company's documentation. The second stage reorganizes this knowledge into a strategy-oriented representation by identifying concepts such as *Company*, *Offering*, *Customer Segment*, *Market Space*, and *Competitor*, together with the semantic relationships connecting them according to the Business Strategy Schema (see top part of Figure 3). The resulting *Business Strategy KG* represents not only what is explicitly stated in the documents, but also how this information should be interpreted within a business strategy perspective. This richer semantic representation constitutes the final output of the proposed methodology and serves as the knowledge base exploited by downstream GraphRAG applications (see bottom part of Figure 3).

# 4. Experimental Evaluation

## 4.1. Experimental Setup

The objective of the experimental evaluation is to assess whether the proposed two-stage methodology improves the usefulness of the *Business Strategy KG* for downstream business question answering. Rather than evaluating the graph solely from a structural perspective, we consider its effectiveness as a retrieval substrate for RAG. This setting reflects the intended use of the generated *Business Strategy KG*, namely supporting language models in answering questions grounded in organizational and strategic knowledge.

For every user question, the retrieval component first identifies the information that is most relevant to answering the query. The retrieved evidence is then provided to the language model together with the original question, allowing the model to generate a grounded answer based on the retrieved context rather than relying exclusively on its internal knowledge. By keeping the LLM, prompting strategy, and answer generation procedure identical across all experiments, differences in performance can be directly attributed to the quality and organization of the retrieved knowledge.

The retrieval process varies according to the evaluated configuration. Depending on the available knowledge source, the retrieval module may search directly within the original company documents, exploit the semantic structure of the intermediate *Document KG*, retrieve information from the *Business Strategy KG* generated during the second stage of the pipeline, or combine multiple knowledge sources. Consequently, the experimental comparison evaluates how progressively richer semantic representations affect the quality of the generated answers. Five retrieval configurations are considered throughout the evaluation:

**Standard RAG**: The retrieval process operates directly over the original company documentation. Relevant textual passages[7] are retrieved from the document collection and provided to the LLM (GPT-5-MINI) as contextual evidence. This configuration represents the traditional RAG baseline and evaluates the benefit of retrieving unstructured textual information.

**Document KG (DKG)**: Instead of retrieving textual passages, this configuration retrieves information from the intermediate *Document KG* produced during the first stage of the proposed methodology. Since the graph organizes entities and relationships extracted from the company documentation into a coherent semantic structure, retrieval exploits explicit connections between concepts rather than isolated text fragments. This configuration evaluates the usefulness of the document-derived graph as an intermediate knowledge representation.

**Business Strategy KG (BSKG)**: Retrieval is performed exclusively over the ontology-guided *Business Strategy KG* generated during the second stage of the pipeline. Unlike the *Document KG*, this representation organizes the available information according to the business strategy ontology, integrating organization-specific evidence with expert-defined strategic concepts. This configuration isolates the contribution of the second stage of the proposed methodology.

**Two-Stage Pipeline (FULL)**: Finally, we evaluate the complete proposed framework by combining the intermediate *Document KG* with the *Business Strategy KG*. The retrieval process can therefore exploit both the factual information extracted directly from company documents and the higher-level semantic organization introduced by the ontology-guided graph construction process. This represents the complete implementation of the proposed methodology.

**LLM only**: This configuration does not employ any retrieval mechanism. Questions are directly submitted to the language model, which generates answers relying exclusively on its parametric knowledge. This baseline measures the contribution of retrieval independently of the generated KG. Similarly to context retrieval in Section 3.2, configurations DKG, BSKG and FULL first retrieve the top-$k$ entry nodes ($k = 3$), then their respective top-$n$ neighboring nodes ($n = 3$). Top-$k$ and -$n$ are ordered by their similarity to the query. The content of these nodes and their relations are formatted in a natural language prompt provided as context to the system. Although the retrieval strategies differ in the knowledge sources they exploit, the remaining components of the question answering pipeline remain

[7] Chunk size = 1024, chunk overlap = 128, embedding model = OpenAI TEXT-EMBEDDING-3-SMALL.

unchanged. In every experiment, the retrieved context is provided to the same language model using an identical prompting strategy, and the generated answers are evaluated using the same automatic metrics adopted in our previous work [28]. This experimental design ensures that observed performance differences can be attributed to the underlying knowledge representation rather than variations in the language model or prompting procedure.

### 4.2. Evaluation Protocol

The objective of the proposed methodology is to construct a *Business Strategy KG* that supports strategic reasoning by integrating organization-specific information with expert-defined business knowledge. Evaluating this objective requires assessing the KG's ability to answer business-oriented questions that extend beyond the information explicitly available in company documents.

To the best of our knowledge, no publicly available benchmark directly targets this setting, namely business-strategy question answering over organization-specific documents and expert-defined strategic concepts. Existing enterprise or business-oriented benchmarks address related problems, such as enterprise knowledge retrieval or financial reasoning, but they do not evaluate the transformation of company-specific evidence into a strategy-oriented knowledge graph. We therefore construct a case-study benchmark tailored to the proposed task. The benchmark was created in collaboration with a domain expert, who manually defined 18 business-oriented questions and their corresponding reference answers. Although small, the benchmark is designed to cover the types of reasoning required in strategic analysis rather than simple fact lookup. These questions require combining organization-specific information extracted from the company documentation with the strategic concepts represented in the *Business Strategy KG*, including competitors, market positioning, customer segments, value propositions, strategic opportunities, and market dynamics. For each question, the evaluated retrieval configuration retrieves the relevant contextual information according to its corresponding retrieval strategy. The retrieved evidence is then provided to the LLM, which generates the final answer. The generated answers are subsequently compared against the expert-written reference answers using the automatic and LLM-based evaluation metrics described in the next section.

## 5. Results

The benchmark comprises 18 questions organized into six categories of increasing inferential demand: factual retrieval (Q1–Q4), synthesis (Q5–Q10), multi-hop reasoning (Q11–Q12), contradiction and tension (Q13–Q14), application of strategy frameworks (Q15–Q16), and cross-domain reasoning (Q17–Q18). The factual category establishes whether a configuration preserves the organizational facts contained in the documents. The remaining categories progressively require a system to combine evidence dispersed across sources, resolve conflicting signals, and apply established analytical frameworks to reach a judgment. This graded structure lets the evaluation distinguish retrieval from the higher-order reasoning that strategic analysis demands, and locate where each configuration's contribution lies.

The reference answers were written by the expert directly from the complete company data room, comprising the case narrative, fourteen executive-interview transcripts, financial and sales exhibits, marketing reports, and an internal strategy document, rather than from any single source. The synthesis and reasoning answers additionally apply recognized strategy frameworks, including the Strategy-in-Action canvas [10], the value disciplines of Treacy and Wiersema [29], and Rumelt's strategy kernel [30]. The reference therefore constitutes a defensible expert baseline grounded in both the organizational evidence and the analytical conventions of the field. The design of the questions also informs how they are scored, based on the metrics proposed in [28]:

- **LLM-as-a-Judge metrics.** An independent LLM evaluator scored each generated answer against the reference answer using a five-point Likert scale (1-5) along four dimensions:
  - Accuracy: The factual correctness of the generated answer relative to the reference answer;

  - Completeness: The extent to which the generated answer covers all the information required to answer the question;
  - Relevance: The extent to which the generated answer directly addresses the question while avoiding unrelated or unnecessary information;
  - Usefulness: The extent to which the answer provides clear, actionable, and practically helpful information to the user.
- **Grounding metrics.** For assessing adherence to the specific phrasing and content of the source business documents we calculated:
  - BERTScore [31] uses deep contextual BERT embeddings to evaluate text similarity in NLP tasks such as summarization, translation, and AI text generation;
  - ROUGE-L [32] gauges the overlap of the longest co-occurring in sequence n-grams between ground-truth and generated texts;
  - BLEU [33, 34] measures the similarity between the system output and the ground-truth using n-gram precision.

For the factual category, the reference answers are close to ground truth and admit near-literal comparison. For the synthesis and reasoning categories, the reference is a strong but not unique answer, since alternative framings can be equally defensible. Accordingly, the qualitatively judged metrics (accuracy, completeness, relevance, and usefulness) are treated as the primary measure for these categories, while the surface-overlap metrics (BLEU, ROUGE-L, and BERTScore) are reported only as secondary indicators, because legitimate paraphrase and differences in framing systematically depress n-gram and embedding overlap. The benchmark also deliberately retains genuine tensions present in the source material, including an unreconciled discrepancy between two reported revenue figures and a gap between the positioning each country head states and the positioning their actions reveal, so that the evaluation rewards configurations that surface and reason about strategic tension rather than those that merely return a single retrieved value.

**Table 2**
Mean $F_1$ values of BERTScore, ROUGE-L, BLEU, and mean values of Accuracy, Completeness, Relevance, Usefulness. Scores are computed on a dataset consisting of 18 questions written by a domain expert.

| Method | BERTScore | ROUGE-L | BLEU | Accuracy | Completeness | Relevance | Usefulness |
|---|---|---|---|---|---|---|---|
| Standard RAG | .8276 | .1452 | **.0132** | 3.778 | 3.278 | 4.389 | 4.167 |
| BSKG | .8163 | .1176 | .0030 | 3.222 | 3.000 | 4.222 | 3.889 |
| DKG | **.8314** | **.1541** | .0105 | 4.000 | 3.667 | 4.500 | 4.167 |
| Full | .8289 | .1497 | .0082 | **4.222** | **3.778** | **4.945** | **4.556** |
| LLM only | .8038 | .0987 | .0045 | 1.444 | 1.556 | 2.222 | 2.500 |

Table 2 reports the results obtained by the five configurations. The full pipeline achieves the highest scores across all four LLM-as-a-Judge dimensions. These results indicate that jointly retrieving from the *Document KG* and the *Business Strategy KG* provides more effective support for strategic question answering than relying on either representation individually. Given the benchmark's limited size, these results should be interpreted as descriptive evidence from a case study rather than as statistically conclusive evidence of performance differences. In particular, small differences in average scores may be affected by individual questions. For this reason, we focus primarily on the overall pattern across retrieval configurations and on the qualitative interpretation of the complementary roles played by the *Document KG* and the *Business Strategy KG*.

The two graph representations contribute complementary forms of knowledge. The *Document KG* preserves detailed organizational evidence and obtains the strongest performance among the individual retrieval sources, whereas the *Business Strategy KG* reorganizes this evidence according to expert-defined strategic concepts. Accordingly, the *Business Strategy KG* alone does not outperform retrieval from the *Document KG*. Its contribution becomes most evident when the two representations are combined:

the document-derived graph supplies fine-grained factual evidence, while the strategy-oriented graph provides higher-level semantic organization and relationships among business concepts.

The results obtained with lexical and semantic similarity metrics present a different pattern. The retrieval performed on the *Document KG* using Microsoft GraphRAG achieves the highest BERTScore and ROUGE-L, while Standard RAG obtains the highest BLEU score. However, the differences among retrieval-based configurations are relatively small, and all BLEU and ROUGE-L scores remain low. This is consistent with the nature of the benchmark: many questions admit multiple defensible answers and require synthesis or interpretation rather than reproducing the wording of the expert reference. Surface-overlap metrics therefore penalize valid paraphrases and alternative strategic framings.

For this reason, the LLM-as-a-Judge dimensions provide the primary evidence for evaluating the usefulness of the generated answers. The substantial advantage of the full pipeline on these dimensions suggests that integrating factual document evidence with the semantic organization of the *Business Strategy KG* is particularly beneficial for questions requiring strategic synthesis and reasoning. The poor performance of the LLM only baseline further confirms that specific organizational knowledge is necessary for answering the benchmark questions reliably.

## 6. Conclusions and Future Work

We presented a two-stage pipeline for constructing Business Strategy Knowledge Graphs from heterogeneous organizational documents. The framework separates factual knowledge extraction from strategic interpretation: the *Document KG* captures evidence from company documents, while the *Business Strategy KG* reorganizes this evidence according to an expert-maintained strategy specification. Representing this specification in Markdown enables the automatic generation of extraction schemas, validation rules, and graph population logic.

Results on a small expert-authored benchmark provide initial evidence that combining the *Document KG* and *Business Strategy KG* improves downstream strategic question answering over using either representation individually. Future work will extend the evaluation to additional organizations, larger benchmarks, graph-centric quality assessment, and human-in-the-loop refinement.

## 7. Limitations

The evaluation is limited to a single industrial case study and a small expert-authored benchmark, reflecting the lack of public benchmarks for business-strategy question answering over organization-specific documents and the difficulty of producing reference answers that require confidential company data and domain expertise. The generated *Business Strategy KG* also depends on the expert-defined conceptual specification. Although the Markdown specification simplifies ontology maintenance and pipeline regeneration, the resulting graph reflects the selected strategic framework and may vary under alternative modeling choices. Finally, we evaluate the graph through downstream question answering rather than intrinsic graph quality. Future work will extend the evaluation to additional organizations and assess graph-centric properties such as completeness, consistency, semantic correctness, and structural quality.

## Declaration on Generative AI

During the preparation of this work, the authors used Grammarly, Gemini Pro in order to rephrase and check grammar and spelling. After using these tools, the authors reviewed and edited the content as needed and take full responsibility for the publication's content.

## Acknowledgments

This work is supported by the Apulia Region through the project SIA LABS funded by the PR FESR-FSE+ 2021-2027 “Applying Generative AI and Knowledge Graphs to Business Strategy” (CUP: B96I25000090007).
Eleonora Ghizzota acknowledges the Ph.D. fellowship within the framework of the Italian Ministerial Decree (D.M.) n. 630, date of D.M.: 24.04.2024 - under the National Recovery and Resilience Plan, Mission 4, Component 2, Investment 3.3 - Ph.D. Project ‘Development and application of Generative Artificial Intelligence models based on the Symbiotic AI paradigm to support the Public Administration’, co-supported by company InnovaPuglia S.p.A. (CUP B91I24000240007).

## References

[1] W. X. Zhao, K. Zhou, J. Li, T. Tang, Z. Dong, Y. Hou, B. Zhang, Y. Min, J. Zhang, P. Liu, et al., A survey of large language models, Frontiers of Computer Science 20 (2026) 2012627.
[2] H. Han, Y. Xie, H. Liu, X. Tang, S. Nag, W. Headden, Y. Li, C. Luo, S. Ji, Q. He, et al., Reasoning with graphs: Structuring implicit knowledge to enhance llms reasoning, in: Findings of the Association for Computational Linguistics: ACL 2025, 2025, pp. 25698–25714.
[3] D. Edge, H. Trinh, N. Cheng, J. Bradley, A. Chao, A. Mody, S. Truitt, D. Metropolitansky, R. O. Ness, J. Larson, From local to global: A graph rag approach to query-focused summarization, arXiv preprint arXiv:2404.16130 (2024).
[4] Y. Hu, Z. Lei, Z. Zhang, B. Pan, C. Ling, L. Zhao, Grag: Graph retrieval-augmented generation, in: Findings of the Association for Computational Linguistics: NAACL 2025, 2025, pp. 4145–4157.
[5] L. Zhong, J. Wu, Q. Li, H. Peng, X. Wu, A comprehensive survey on automatic knowledge graph construction, ACM Comput. Surv. 56 (2023). URL: https://doi.org/10.1145/3618295. doi:10.1145/3618295.
[6] M. Hofer, D. Obraczka, A. Saeedi, H. Köpcke, E. Rahm, Construction of knowledge graphs: Current state and challenges, Information 15 (2024) 509.
[7] H. Mintzberg, B. Ahlstrand, J. Lampel, Strategy safari, volume 2, Free press New York, 1998.
[8] M. E. Porter, How competitive forces shape strategy, in: Readings in strategic management, Springer, 1979, pp. 133–143.
[9] K. Vigfússon, L. Jóhannsdóttir, S. Ólafsson, Obstacles to strategy implementation and success factors: A review of empirical literature, Strategic Management-International Journal of Strategic Management and Decision Support Systems in Strategic Management 26 (2021).
[10] G. SCARSO-BORIOLI, D. Sola, R. QUAGLIA, A. Jordan, Towards a machine-readable standard ontology for business strategy, ESCP Business School Impact Paper (2024).
[11] T. Brown, B. Mann, N. Ryder, M. Subbiah, J. D. Kaplan, P. Dhariwal, A. Neelakantan, P. Shyam, G. Sastry, A. Askell, et al., Language models are few-shot learners, Advances in neural information processing systems 33 (2020) 1877–1901.
[12] S. Kim, J. Ha, H. Lee, S. Park, S. Cho, Human-guided collective llm intelligence for strategic planning via two-stage information retrieval, Information Processing & Management 63 (2026) 104288.
[13] M. J. Saeedizade, E. Blomqvist, Navigating ontology development with large language models, in: European semantic web conference, Springer, 2024, pp. 143–161.
[14] A. S. Lippolis, M. J. Saeedizade, R. Keskisärkkä, S. Zuppiroli, M. Ceriani, A. Gangemi, E. Blomqvist, A. G. Nuzzolese, Ontology generation using large language models, in: European Semantic Web Conference, Springer, 2025, pp. 321–341.
[15] A. Oyewale, T. Soru, Llm-driven ontology construction for enterprise knowledge graphs, arXiv preprint arXiv:2602.01276 (2026).
[16] M. S. Abolhasani, R. Pan, Leveraging llm for automated ontology extraction and knowledge graph generation, arXiv preprint arXiv:2412.00608 (2024).

[17] G. Ouyang, Y. Tang, Y. Huang, Llm4onto: Automated domain ontology construction using large language models (2016).
[18] V. K. Kommineni, B. König-Ries, S. Samuel, From human experts to machines: An llm supported approach to ontology and knowledge graph construction, arXiv preprint arXiv:2403.08345 (2024).
[19] B. Zhang, H. Soh, Extract, define, canonicalize: An llm-based framework for knowledge graph construction, in: Proceedings of the 2024 conference on empirical methods in natural language processing, 2024, pp. 9820–9836.
[20] H. Bian, Llm-empowered knowledge graph construction: A survey, arXiv preprint arXiv:2510.20345 (2025).
[21] L. Cotti, I. Drago, A. Rula, D. Bianchini, F. Cerutti, Ontologx: Ontology-guided knowledge graph extraction from cybersecurity logs with large language models, Advanced Intelligent Systems (2026) e202501381.
[22] S. Khorshidi, A. Nikfarjam, S. Shankar, Y. Sang, Y. Govind, H. Jang, A. Kasgari, A. McClimans, M. Soliman, V. Konda, et al., Odke+: Ontology-guided open-domain knowledge extraction with llms, arXiv preprint arXiv:2509.04696 (2025).
[23] P. Lewis, E. Perez, A. Piktus, F. Petroni, V. Karpukhin, N. Goyal, H. Küttler, M. Lewis, W.-t. Yih, T. Rocktäschel, et al., Retrieval-augmented generation for knowledge-intensive nlp tasks, Advances in neural information processing systems 33 (2020) 9459–9474.
[24] B. Peng, Y. Zhu, Y. Liu, X. Bo, H. Shi, C. Hong, Y. Zhang, S. Tang, Graph retrieval-augmented generation: A survey, ACM Transactions on Information Systems 44 (2025) 1–52.
[25] Q. Zhang, S. Chen, Y. Bei, Z. Yuan, H. Zhou, Z. Hong, H. Chen, Y. Xiao, C. Zhou, J. Dong, et al., A survey of graph retrieval-augmented generation for customized large language models, arXiv preprint arXiv:2501.13958 (2025).
[26] V. A. Traag, L. Waltman, N. J. Van Eck, From louvain to leiden: guaranteeing well-connected communities, Scientific reports 9 (2019) 1–12.
[27] J. J. Miller, Graph database applications and concepts with neo4j, in: Proceedings of the southern association for information systems conference, Atlanta, GA, USA, volume 2324, AISEL, 2013, pp. 141–147.
[28] E. Ghizzota, A. Jordan, A. Petruzzelli, L. Siciliani, G. Spillo, P. Basile, D. Sola, G. Scarso Borioli, G. Semeraro, Integrating knowledge graph and large language models for defining business strategies, in: G. Sérasset, K. Gkirtzou, M. Cochez, J.-C. Kalo (Eds.), Proceedings of the Knowledge Graphs and Large Language Models Workshop (KG-LLM) @ LREC26, European Language Resources Association (ELRA), Palma, Mallorca, Spain, 2026, pp. 53–62. doi:10.63317/4e642nbxhuqr.
[29] M. Treacy, F. Wiersema, Customer intimacy and other value disciplines, Harvard Business Review 71 (1993) 84–93.
[30] R. P. Rumelt, Good Strategy Bad Strategy: The Difference and Why It Matters, Crown Business, 2011.
[31] T. Zhang, V. Kishore, F. Wu, K. Q. Weinberger, Y. Artzi, Bertscore: Evaluating text generation with BERT, CoRR abs/1904.09675 (2019). URL: http://arxiv.org/abs/1904.09675. arXiv:1904.09675.
[32] C.-Y. Lin, Rouge: A package for automatic evaluation of summaries, in: Text summarization branches out, 2004, pp. 74–81.
[33] M. Post, A call for clarity in reporting BLEU scores, in: O. Bojar, R. Chatterjee, C. Federmann, M. Fishel, Y. Graham, B. Haddow, M. Huck, A. J. Yepes, P. Koehn, C. Monz, M. Negri, A. Névéol, M. Neves, M. Post, L. Specia, M. Turchi, K. Verspoor (Eds.), Proceedings of the Third Conference on Machine Translation: Research Papers, Association for Computational Linguistics, Brussels, Belgium, 2018, pp. 186–191. URL: https://aclanthology.org/W18-6319/. doi:10.18653/v1/W18-6319.
[34] K. Papineni, S. Roukos, T. Ward, W. jing Zhu, Bleu: a method for automatic evaluation of machine translation, 2002, pp. 311–318.