# Chapter 10 SPG Embracing the New Era of Cognitive Intelligence

The future trend of enterprise digitalization and intelligent upgrading is to build knowledge based on the massive business data accumulated in the enterprise's big data system, promoting data knowledgeization. By integrating business data with AI systems, business intelligence can be achieved. Property graphs have the advantage of compatibility with big data systems, and SPG, based on property graph, aims to accelerate the data knowledgeization and the organic integration of knowledge and AI systems. This chapter combines the core capabilities and two case studies introduced in previous chapters to summarize and analyze the strengths, limitations, opportunities, and challenges of SPG.

## 10.1 SWOT Analysis of SPG Compared to Property Graphs

Based on the previous descriptions, we can analyze the strengths, weaknesses, opportunities, and threats of SPG from the four quadrants of SWOT.

- **Strengths of SPG.** (1) SPG has a low cost and is compatible with big data architecture, allowing for the rapid construction of domain knowledge graphs based on structured data accumulated in enterprise-level applications. (2) The hierarchical semantic model of SPG supports the continuous evolution of incomplete knowledge graphs, meeting the requirements of rapid business deployment, continuous data accumulation and improvement, and gradual technological application in industrial applications. (3) SPG overcomes the semantic limitations of LPG and effectively connects big data with LPG node/edge structures, enhancing semantic connections and better integrating AI technologies.

- **Weaknesses of SPG.** SPG is still in the growth stage, and there are some compromises and weaknesses in the design of its capabilities, which we need to continuously overcome in the later stage. (1) How does dynamic classification achieve inheritance and extension? Currently, the dynamic classification model effectively addresses the granularity issue of types but has limitations in application, making it difficult to extend properties under new subclasses. (2) Continuous improvement of the built-in semantic structures of the entity is necessary. The current built-in semantic structures in the entity model are not sufficiently enriched, with definitions and constraints limited to event subjects/objects/times and hierarchical concepts. To meet the demands of controlled generation and interpretable reasoning, clear expression of built-in semantic structures within entities requires ongoing improvement in collaboration with downstream applications. (3) Instance-concept linkage reasoning model. While SPG possesses some inductive reasoning capabilities from instances to concepts, there is still significant room for improvement in the co-conduction of concepts and instances and the deductive reasoning from concepts to instances. This requires continuous optimization through the application and refinement of various causal knowledge graphs.

SPG provides a scalable, AI-integrated knowledge graph system but still needs improvements in dynamic classification, semantic structures, and instance reasoning.

- **Opportunities for SPG.** (1) Filling the gap in semantic frameworks for enterprise knowledge graph applications. RDF/OWL, due to their complexity, have not effectively landed in enterprises. Establishing enterprise-level standardization to facilitate cross-entity knowledge semantic alignment, and to promote the circulation, interoperability, exchange, and sharing of knowledge more conveniently. (2) Driving the development of a universal engine architecture for building knowledge graphs. This promotes the democratization and accessibility of the knowledge graph technology. Large-scale applications in various technical fields rely on standardization and framework development, as seen in search engines, deep learning, and cloud computing. (3) Bridging the gap between knowledge graphs and LLMs. Enterprises can quickly incubate/fine-tune new pre-training models based on Transformer or open-source LLMs. Using the standardized symbols of SPG, efficient knowledge injection, associative prompts, knowledge queries, and other tasks can be achieved during the pre-training, SFT/RLHF, and inference stages, and forming a stable paradigm for the interaction between knowledge graphs and LLMs. Additionally, by data knowledgeization, we aim to construct a symbolic world domain knowledge system that complements and is equivalent to the neural network-based LLM knowledge system.

- **Challenges for SPG.** (1) Performance challenges in scaling applications, particularly during the knowledge construction phase. The performance overhead of extraction models and entity linking can significantly impact the efficiency of large-scale knowledge graph construction. (2) Further refinement of system capabilities. The capabilities of the SPG system need continuous optimization in conjunction with more business scenario and applications. (3) Cultivating semantic understanding in users' mental models. On one hand, there is a need to continuously improve users' understanding of semantics, and on the other hand, efforts should be made to reduce users' perception of semantics.

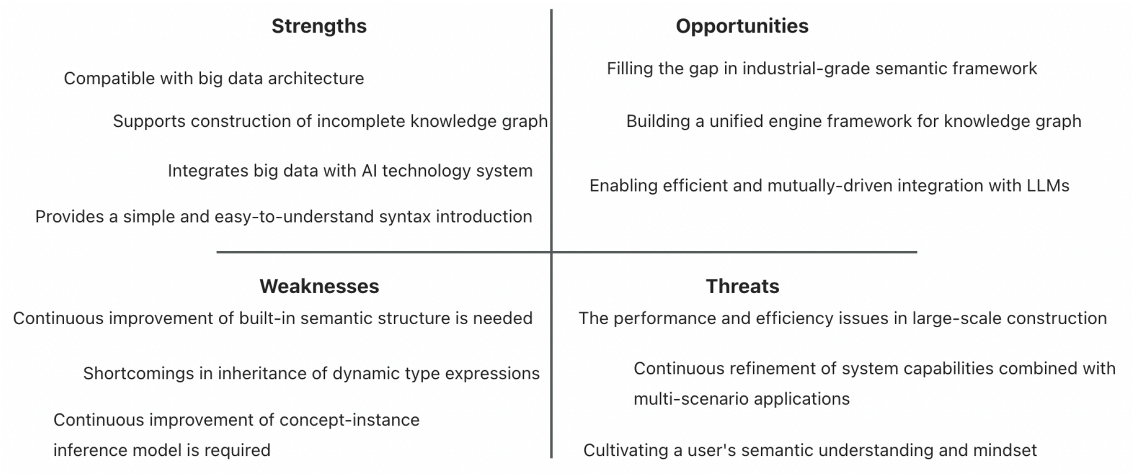


Figure 55: SWOT Analysis of SPG

SPG fills critical gaps in enterprise AI knowledge graphs but must overcome scalability, optimization, and user adoption challenges for broader implementation.

# 10.2 Problem Resolution and Outstanding Issues from Chapter 2

Table 15: Fundamental Problems of Knowledge Management Based on LPG and the Resolution Status by SPG

| Typical Problems of LPG | | SPG Solutions | SPG Status |
|---|---|---|---|
| Type Management | Semantic deficiency in subjects | Event, entity, and concept classification | Supported |
| | Type granularity issue | Dynamic types | Supported |
| | Difficulty in property/relation selection | Standard properties, concept types | Supported |
| | Entities of the same type with different names | Knowledge fusion | Supported |
| Incompleteness Knowledge Graph Construction | Entity construction from multiple data sources | Entity linking, entity resolution | Partial support |
| | Knowledge reuse across multiple knowledge graphs | Knowledge fusion | Supported |
| Logical Dependencies | Semantic deficiency in logical predicates | Semantic predicates with rule conditions | Partial support |
| | Inconsistency due to separation of definition and logic | Logical Derivation | Supported |
| | Window-type property/relation explosion | Window standard types, variable parameter properties | Supported |
| | Inconsistent property logic | Logical properties | Supported |
| | Inconsistent relation logic | Logical relations | Supported |
| Logical reasoning of causal | Obstruction in event logic propagation | Activating new event instances during causal deduction | Supported |
| | Event Classification | Dynamic types | Supported |
| | Semantic Composition of Concepts | Dynamic composition based on event and entity facts | Supported |

It is important to note that Table 15 primarily lists the fundamental problems of knowledge management based on LPG and the resolution status by SPG. It mainly focuses on the semantic aspects of entities and logical predicate semantics. The programmable framework and complex knowledge reasoning are built upon a knowledge management framework in a virtuous cycle. They are not included in the basic capabilities of knowledge management and are not listed in this table. However, they will be further described in the future release plan in Chapter 11.

SPG resolves key issues in knowledge management from LPG, particularly in semantic representation, logical reasoning, and event classification, while outlining fundamental improvements over LPG.

# Chapter 11 Outlook on the Future of SPG

This whitepaper has addressed the challenges faced by enterprise-level knowledge management and discussed the higher requirements for knowledge semantic representation and engine frameworks due to changes in demand paradigms in enterprise-level knowledge graph applications. In Chapter 1, we summarized some of the key problems that still exist in the development of knowledge graph technology:

- Lack of unified semantic representation. Currently, strong semantic knowledge graphs have not achieved industrial implementation based on RDF/OWL, while weak semantic property graphs (LPG) are widely used in industrial-grade knowledge graphs.
- Multiple tools but lack of standardization. The development of customized extraction algorithms/entity linking algorithms for each data set, graph database-backed graph storage, representation learning tools, fuzzy retrieval tools, knowledge query tools, and other tools have led to significant dispersion and inconvenience in the application of knowledge graph technology.

In order to achieve large-scale industrial application of any complex technology, it is necessary to have a unified technical framework that shields complex technical details and supports rapid deployment of new businesses. It also requires a modular architecture that allows for layering and decoupling of domain models and core engines, enabling fast migration to new domains. The same applies to knowledge graphs. The development of knowledge graph technology needs to keep up with the times. SPG defines an industrial-grade, user-friendly knowledge semantic framework for strong semantic knowledge graphs. It helps enterprises accelerate the knowledgeization of massive amounts of data. Through the unified technical framework and engine architecture provided by the SPG knowledge engine, the technology can be standardized, democratized, and made accessible to a wide range of users.

Looking towards the future, knowledge graphs have vast application potential. On one hand, as the best modeling practice for structured data, knowledge graphs can unify data modeling from various perspectives such as machines, algorithms, engineering, business, and operations. They can build next-generation data architectures in line with the concept of data fabric, accelerating the knowledgeization of massive enterprise data, connecting data silos, discovering implicit relations, unlocking the full value of data, and reducing the cost of finding and using data, ultimately bringing greater growth opportunities for businesses. On the other hand, knowledge graphs complement LLMs perfectly. Knowledge graphs have characteristics such as strong facts, weak generalization, strong interpretability, low computational cost, and high construction cost. In contrast, LLMs have weak facts, strong generalization, poor interpretability, high computational cost, and strong semantic understanding. In the future, the goal is to achieve efficient cooperation and complementarity between unified knowledge symbol representation and engine architecture and LLMs. Through the advancement of LLM technology, the cost of knowledge graph construction can be further reduced, accelerating data knowledgeization and providing additional domain knowledge for controlled generation based on LLMs. The construction of massive common-sense domain knowledge bases can also accelerate the progress of general artificial intelligence. Realizing the integration and complementarity of

The future of SPG focuses on addressing current challenges like the lack of semantic standardization and tool fragmentation while aiming for tighter integration with LLMs and scalable industrial applications.

the knowledge graphs and LLMs strongly relies on a comprehensive knowledge graph and LLM technology stack. Currently, LLM technology has matured, and with the strong semantic knowledge graph framework defined by SPG, it is expected to form a seamless application framework that can be seamlessly integrated with LLMs. This framework will enable industrial-level, generalizable, highly robust, and interpretable comprehensive artificial intelligence technologies based on knowledge graphs and LLMs.

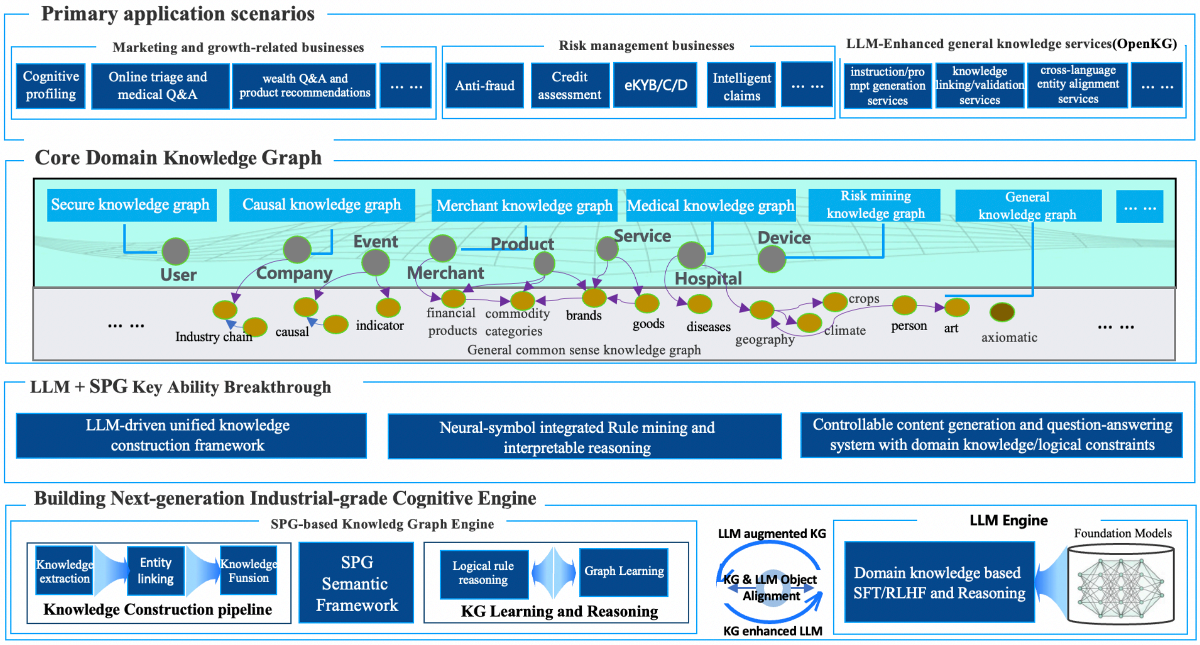


Figure 56: Future Outlook on the Dual-Drive Technical Paradigm of SPG and LLM

The fusion of symbolic logic and neural networks has always been a research hotspot in the industry. One common approach is to use neural networks to learn the rules and relations in symbolic logic, enabling them to better handle complex logical problems. Another approach is to use symbolic logic to guide the learning process of neural networks, improving their accuracy and interpretability. Knowledge graphs, as a typical representative of symbolic logic, have unique advantages in structural representation, semantic characterization, and knowledge association. The unified semantic framework provided by SPG can provide them with stronger vitality. Currently, the fusion of neural networks and symbolic logic mainly occurs in the knowledge reasoning stage. With the emergence of LLMs, new ideas are provided for the integration of symbolic logic and neural networks. On one hand, knowledge graphs, as the underlying support for semantic representation of symbolic logic and knowledge data management, can leverage the powerful semantic understanding capabilities of LLMs and the strong structure and semantics of knowledge graphs to automate prompts and sample construction. This can help the knowledge graphs form a unified knowledge extraction framework and accelerate the knowledgeization of data. On the other hand, in the content generation stage, applying domain knowledge data with strong semantic constraints can effectively avoid the problem of hallucinations and nonsense in LLMs. These issues are expected to be accelerated addressed in the SPG + LLM paradigm. We will continue to improve the expressive capabilities of SPG through industrial practices and enhance LLM through SPG to achieve alignment with objective facts, effectively avoiding/reducing model hallucinations. At the same time, LLM will also enhance SPG to improve the conversion efficiency

SPG aims to integrate with LLMs to create an advanced AI framework, leveraging symbolic logic and structured knowledge graphs to improve accuracy, reasoning, and mitigate hallucinations in LLMs.

of data knowledgeization. We are committed to building a next-generation artificial intelligence engine driven and enhanced by both SPG and LLM.

Table 16: Future Release Plan for SPG

| Stage | Release Content | Release Sub-items |
|---|---|---|
| Stage 1 | "Semantic-Enhanced Programmable Knowledge Graph (SPG) White Paper" V1.0 ---- Next-generation industrial-grade knowledge semantic framework.<br>Introduce the background, specifications, framework, and case studies of SPG knowledge semantic framework. | 1.Overall framework of SPG<br>2.Two case studies with rich content<br>3.SPG-Schema Core<br>4.Basic requirements of SPG-Engine<br>5.Basic requirements of SPG-controller<br>6.Basic framework of SPG-Programming |
| | Establish SPG community and welcome its establishment | 1.Establishment of community operation mechanism |
| | Open-source SPG Engine v1.0 | 1.Build an open-source engine that meets the requirements of White Paper 1.0<br>2.Create an open-source community based on the engine<br>3.Default adaptation to Tugraph technology stack, support vendor extensions |
| | Recommendation of standardization for SPG [Industry, National, International] | 1.Contact various standardization organizations to promote standardization projects |
| Stage 2 | "Semantic-Enhanced Programmable Knowledge Graph (SPG) White Paper" V2.0 ---- Next-generation industrial-grade knowledge semantic framework.<br>Introduce SPG programmable framework Operator, SDK, case studies, etc., open up a complete knowledge construction framework. | 1.Improved SPG-schema<br>2.Complete definition of operators and framework in SPG programmable framework<br>3.Support mapping-based [programmable] knowledge graph construction, support search engines, vector retrieval, model services, etc.<br>4.Basic inference solution, including basic query language for construction, query, and inference<br>5.Complete storage solution [standalone]<br>6.Basic requirements of SPG+LLM and exploration in knowledge extraction/NL2KGDSL |
| | Release a query language that meets semantic and programmable requirements | 1.Release GQL or KGDSL that meets the basic requirements of SPG applications<br>2.CRUD operations in SPG<br>3.Support basic natural language queries |
| | Open-source framework V2.0 | 1.Complete open-source implementation of SPG framework including White Paper 2.0<br>2.Query language implementation module [GQL or KGDSL]<br>3.Programmable SDK |
| | Publish SPG industry and international standards | 1.Finalize the standard draft<br>2.Publicize the standard draft |
| Stage 3 | "Semantic-Enhanced Programmable Knowledge Graph (SPG) White Paper" V3.0 ---- Next-generation industrial-grade knowledge semantic framework.<br>Introduce SPG inference engine, graph learning, rule-guided interpretable inference, etc., open up complete knowledge inference capabilities. | 1.Programmable module to accumulate reusable model hubs, layer hubs, etc., supporting out-of-the-box use<br>2.SPG industry extensions for domains such as space-time, healthcare, etc.<br>3.Comprehensive syntax set for rule-based inference engine<br>4.SPG and AGL linkage to support knowledge inference, interpretable inference, etc. |
| | Open-source framework V3.0 | 1.Fully match the final standard draft and SPG V3.0<br>2.Extension modules |
| | Knowledge construction and open-source solutions based on common sense knowledge using SPG | Build an open-source, crowdsourced, service community for the healthy growth of common sense knowledge based on OpenKG |
| Stage 4 | "Semantic-Enhanced Programmable Knowledge Graph (SPG) White Paper" V4.0 ---- Next-generation industrial-grade knowledge semantic framework.<br>Introduce the paradigm of SPG + LLM double-driven and construct a controllable LLM technology system. | 1.Complete paradigm and system framework of SPG + LLM<br>2.Basic natural language interface for construction, query, and inference<br>3.Natural language interaction module [based on LLM]<br>4.Supports extractive construction (based on LLM) |
| | Knowledge services for common sense knowledge based on SPG. | Build an enhanced knowledge services with LLMs, such as instruction/prompt generation services, entity linking/validation services, etc., to promote knowledge element exchange and growth. |

In the future, we will continue to upgrade SPG. Table 16 represents our planned release content, and the release timeline will be updated on the SPG official account: "Semantic-Enhanced Programmable Knowledge Graph Framework". We welcome your attention and interaction, and together, we can explore the industrial-grade knowledge graph architecture paradigm.

The SPG release roadmap outlines the stages of implementation, including open-source expansions, query language improvements, and industry adoption, ensuring the evolution of an industrial-grade knowledge graph framework.

语义增强可编程图谱框架

微信扫描二维码，关注我的公众号