The Maersk Mc-Kinney Moller Institute

EMNLP2026 BUDAPEST

AUTOMATED KNOWLEDGE BASE CONSTRUCTION 2026 | BUDAPEST

SDU University of Southern Denmark

Edge Knowledge Bases for Sustainable and Explainable AI

Aparna S. Varde[1], Emil Njor[1], Francesco Daghero[2], Mahyar Tourchi Moghaddam[2]

[1]Center for Software Technology, SDU Vejle | [2]Dept. of Software Engineering, SDU Odense

Maersk Mc-Kinney Moller Institute (MMMI), University of Southern Denmark (SDU), Denmark

Introduction

We propose a vision of **EdgeKBs** (Edge Knowledge Bases) integrating KBs into Edge AI systems.

EdgeKBs can infuse commonsense and domain knowledge within Edge devices to guide them while learning from training data.

By doing so, we not only bring explainability and performance benefits of KBs to the Edge, but also enable more efficient, sustainable Edge AI systems.

We envisage three approaches to develop EdgeKBs: Slicing \& Dicing; Tailoring to Fit; and Dynamic EdgeKBs. These

EdgeKBs will bring Knowledge-Guided Machine Learning (KGML) to the Edge to support next-generation applications, e.g., Physical AI, cobots, and local natural language interfaces.

Background

Deploying AI at the network edge delivers lower latency, higher bandwidth efficiency, and enhanced privacy, but is constrained by strict hardware limits in memory, compute, and energy. Knowledge-Guided Machine Learning (KGML) integrates structured knowledge into machine learning models to boost explainability and generalizability. Introducing EdgeKBs (Edge Knowledge Bases) brings structured domain knowledge directly to resource-constrained devices, enabling high predictive performance with smaller, energy-efficient base models.

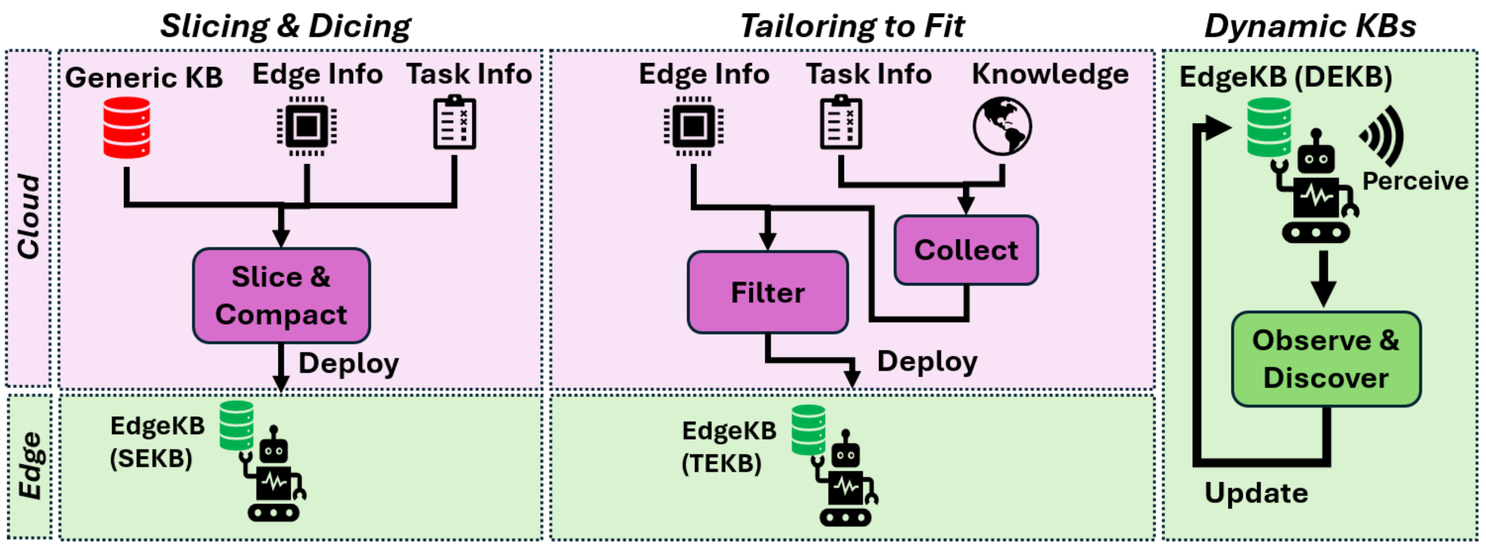
Key Applications

Edge Natural Language Interfaces: Powers offline, grounded conversational agents for critical industrial workflows.

Private Smart Assistants: Merges local commonsense with user routines on-device for complete privacy.

Physical AI & Cobots: Supplies cobots with real-time affordances, task constraints, and safety rules.





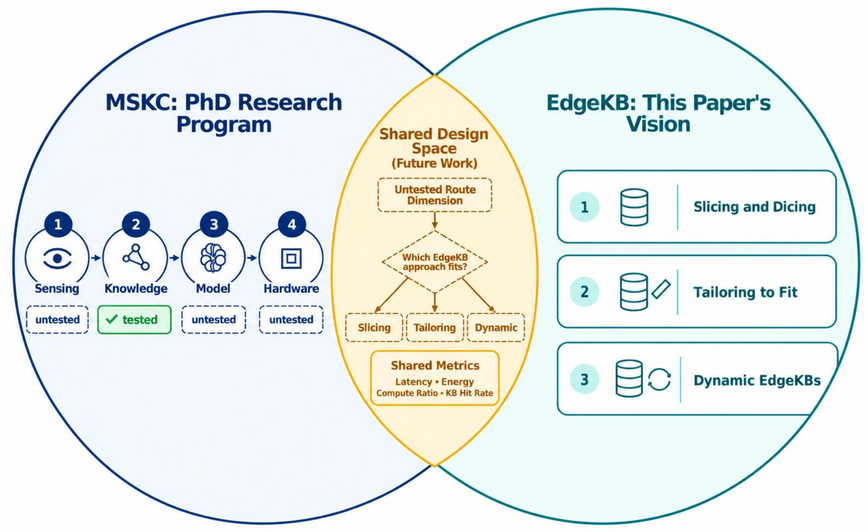
Proposed EdgeKB Approaches

We propose three archetype approaches to construct EdgeKBs.

Slicing and Dicing (SEKB): Extracts targeted sub-graphs from large-scale KBs using multifaceted scoring (plausibility, typicality, salience) or LLM extraction (semantic chunking and SPO triples).

Tailoring to Fit (TEKB): Constructs customized KBs from scratch to match application-specific semantics, custom data structures, and specialized edge hardware.

Dynamic EdgeKBs (DEKB): Allows on-device KBs to continuously evolve post-deployment via Small Language Models (SLMs) through Knowledge Graph refinement or autonomous entity/relation discovery.



Experimentation Example — MSKC (minimum-sufficient knowledge computation) in targeted applications, e.g., decision support of wind turbines for renewable energy, in conjunction with EdgeKB development. Knowledge routing is partly tested, the rest of the parts can be mounted with EdgeKB components.

Table 1: Evaluation Metrics for Edge Knowledge-Guided Machine Learning Systems

Category	Metric	Description	Unit
Hardware	Static Memory (Storage)	Disk/Flash space required to store AI model parameters and the KB index.	MB / GB
	Dynamic Memory (RAM)	Peak working memory consumption during concurrent KB retrieval and model inference.	MB / GB
	Energy Consumption	Total energy consumed per query/batch, critical for battery-constrained edge devices.	mJ / J
	Latency	End-to-end processing time from input reception to final output generation.	ms
AI Model	Accuracy	Proportion of correct predictions or classifications out of all evaluated samples.	%
	F1-Score / Precision	Performance metrics for robustness, especially useful for imbalanced edge datasets.	0–1
Unified	Storage Ratio (KB:AI)	Ratio of static memory footprint of the KB versus the AI Model parameters.	Ratio
	Compute Ratio (KB:AI)	Proportion of processing time (or FLOPs) spent on KB retrieval versus AI model inference.	Ratio
	KB Hit/Relevance Rate	Percentage of inferences where the KB successfully provided useful context.	%

Evaluation

Integrating EdgeKBs into local AI frameworks establishes a foundation for sustainable EdgeKGML. Grounding local inference in explicit knowledge allows systems to achieve high predictive accuracy while utilizing significantly smaller, energy-efficient base models.

Evaluating these systems requires balancing standard hardware and AI performance metrics alongside unified cross-domain metrics.

Hardware: Static Memory (Flash), Dynamic RAM, Energy per Query (mJ/J), and Latency (ms).

AI Performance: Accuracy, Precision, and F1-Score.

Unified System Metrics: Storage Ratio (KB vs. AI model footprint), Compute Ratio (retrieval vs. inference time), and KB Hit/Relevance Rate.

Acknowledgements

Aparna Varde acknowledges a research grant NF25OC011003 from NNF, Denmark.

Francesco Daghero and Mahyar Tourchi Moghaddam acknowledge the REINS project, supported by Next-Generation AI (Forskningsreserven) and Digital Research Centre Denmark (DIREC).

We thank Purbak Sengupta, a PhD candidate at CST, MMMI, SDU Vejle, for his inputs related to an illustration in this paper.

All the authors are at the Maersk Mc-Kinney Moller Institute (MMMI) in the Faculty of Engineering (TEK) at the University of Southern Denmark (SDU).

sdu.dk