Semantic Technology in Your Pocket

Peter Crocker[1], Ian Horrocks[1] and Yavor Nenov[1]

[1]Oxford Semantic Technologies, Oxford, UK

Abstract

In January 2025, Samsung announced the launch of the Personal Data Engine (PDE) on their flagship Galaxy S25 smartphone. The PDE uses semantic technology in the form of the RDFox® Knowledge Graph system to provide on-device AI capabilities to client applications. This almost certainly represents the largest ever deployment of semantic technology, with millions of users now carrying a semantic reasoning engine in their pocket.

Keywords

Knowledge Graphs, Semantic Technology, Rules-based AI

1. Introduction

In January 2025, Samsung announced the launch of the Personal Data Engine (PDE) on their flagship Galaxy S25 smartphone [1]. The PDE uses a Knowledge Graph (KG) to integrate personal data derived from multiple sources and to provide on-device AI capabilities to client applications. The KG system is an Android build of the RDFox® system from Oxford Semantic Technologies. RDFox® stores data in the form of an RDF graph and an ontology in the form of OWL RL axioms and Datalog rules; it provides a high-performance incremental reasoning engine and a SPARQL query interface. This almost certainly represents the largest ever deployment of semantic technology, with millions of users now carrying a semantic reasoning engine in their pocket. In the remainder of this paper we will briefly explain what the PDE is, how it uses semantic technology and how RDFox® can enable this on a mobile device.

2. Personal Data Engine

The PDE collects data from a wide range of sources including multiple apps and system software (see Figure 1). This can include everything from structured data through to text and (possibly moving) images. The PDE analyses this data using appropriate AI technologies including machine learning and LLMs. The result of this analysis is stored in the user's personal knowledge graph where it is augmented with triples entailed by the ontology. Over time, the knowledge graph builds up a holistic view of the user and their surrounding context. This is used to provide services to client applications supporting features such as personalisation and recommendation. For example, the PDE supports the Now Brief function on the S25 which provides a personalised briefing of selected content that updates throughout the day. By running all of this capability encrypted on the device users can be assured by the highest levels of privacy.

The PDE uses an android build of the RDFox® system to store, reason over and query the graph. Importantly, all this happens on the device. This helps to guarantee privacy by avoiding personal data being moved off the device.

3. RDFox®

To support AI features in client applications the PDE must provide real-time answers to complex queries over the KG and ontology. To achieve this, RDFox® uses patented in-memory data structures and a

ISWC 2025 Companion Volume, November 2–6, 2025, Nara, Japan

peter.crocker@oxfordsemantic.tech (P. Crocker); ian.horrocks@oxfordsemantic.tech (I. Horrocks); yavor.nenov@oxfordsemantic.tech (Y. Nenov)

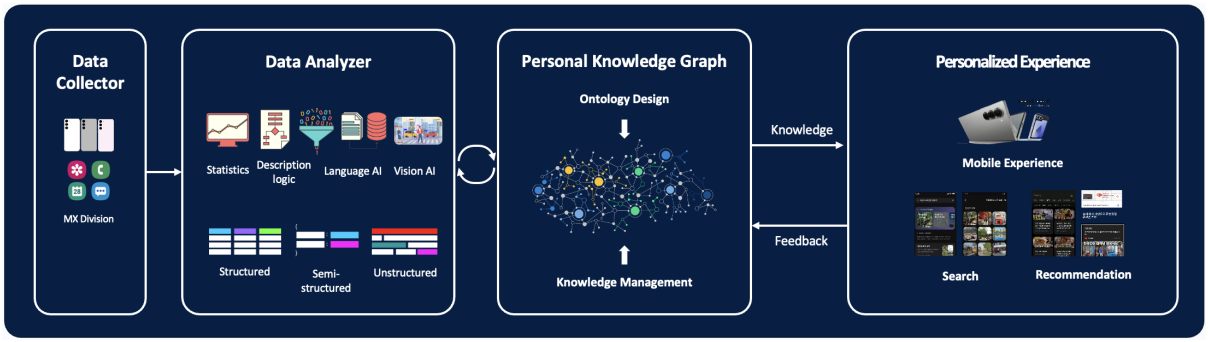

Figure 1: The Personal Data Engine

highly optimised incremental reasoning engine [2]. Figure 2 illustrates the basic architecture of the system. The reasoning engine exploits modern multi-core architecture to parallelise reasoning and incrementally maintains a fully materialised graph that extends the RDF data with triples entailed by the ontology [3]. A highly optimised SPARQL query engine provides access to the materialised graph.

Materialisation effectively means repeatedly answering queries corresponding to the bodies of axioms/rules and adding the answer triples back into the graph [4]. An RDF graph is just a set of subject, predicate, object triples, and query answering involves computing joins over the triples. To do this efficiently, we need the triples to be accessible in various orders, for example in subject order or object order, and we also need indices that allow us to rapidly find appropriate entry points into the various orderings. Moreover, we need to update the relevant data structures as new triples are added, and we need to do this with minimal locking so as to allow for efficient parallelisation. These considerations motivated the development of novel "lock-free" data structures for RDFox® [5].

In settings such as the PDE, where data is regularly changing, incremental maintenance of the materialisation is essential. RDFox® implements incremental addition and retraction using a novel extension of the Delete rederive algorithm (DRED) [6]. This works by first eagerly over-deleting all facts that depend (possibly indirectly) on a deleted fact, and then rederiving facts that still hold due to some alternative derivation. This has the advantage that it doesn't require any additional data structures, and it can handle arbitrary rules, including recursive rules; however, eager over-deletion and the subsequent rederivation of facts can be very inefficient, particularly when facts participate in long inference chains. In RDFox®, deletion is interleaved with checks to determine if facts still hold via alternative derivations, which avoids excessive over-deletion and hence improves the efficiency of incremental materialisation.

Finally, SPARQL is a complex language, and even with all entailed facts materialised efficient query answering can still be challenging. RDFox® includes a highly optimised SPARQL query engine with (patented) sideways information passing [7].

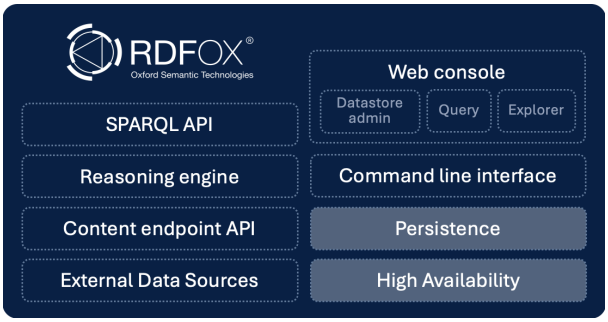

Figure 2: RDFox® System Architecture

4. Future Vision

Development of the PDE was carried out in 2024 and the PDE was deployed on Samsung's flagship S25 smartphone platform in 2025. However, this is only the beginning—with Samsung's other 'phones and their position in the consumer electronic devices market, the opportunities for deployment of the PDE are vast. With RDFox® powering the PDE, Samsung's comprehensive ecosystem of devices puts them in a unique position to provide a holistic and deeply personalised user experience not only across multiple apps but also across a wide range of different devices.

Declaration on Generative AI

The authors have not employed any Generative AI tools.

References

[1] S. Electronics, Samsung Galaxy Unpacked January 2025, Jan. 2025. URL: https://www.youtube.com/live/HinL5jCy_oI.

[2] B. Motik, Y. Nenov, I. Horrocks, Parallel materialisation of a set of logical rules on a logical database, U.S. Patent 10817467, Oct. 2020.

[3] B. Motik, Y. Nenov, R. Piro, I. Horrocks, D. Olteanu, Parallel materialisation of Datalog programs in centralised, main-memory RDF systems, in: Proc. of the 28th Nat. Conf. on Artificial Intelligence (AAAI 14), AAAI Press, 2014, pp. 129–137.

[4] S. Abiteboul, R. Hull, V. Vianu, Foundations of Databases, Addison Wesley Publ. Co., Reading, Massachussetts, 1995.

[5] Y. Nenov, R. Piro, B. Motik, I. Horrocks, Z. Wu, J. Banerjee, RDFox: A highly-scalable RDF store, in: Proc. of the 14th International Semantic Web Conference (ISWC 2015), Lecture Notes in Computer Science, Springer, 2015.

[6] B. Motik, Y. Nenov, R. Piro, I. Horrocks, Maintenance of datalog materialisations revisited, Artificial Intelligence 269 (2019) 76–136.

[7] B. Motik, Y. Nenov, I. Horrocks, Complex query evaluation using sideways information passing, U.S. Patent 11216456, Jan. 2022.