

AN ECOSYSTEM FOR PERSONAL KNOWLEDGE GRAPHS



We propose an ecosystem for personal knowledge graphs (PKG), commonly defined as resources of structured information about entities related to an individual, their attributes, and the relations between them. PKGs are a key enabler of secure and sophisticated personal data management and personalized services. We propose our own definition of a PKG, emphasizing the aspects of (1) data ownership by a single individual and (2) the delivery of personalized services as the primary purpose. We argue that a holistic view of PKGs is needed to unlock their full potential and propose a unified framework for PKGs, where the PKG is part of a larger ecosystem with clear interfaces toward data services and data sources.

AUTHORS

Martin G. Skjæveland, Krisztian Balog, Nolwenn Bernard, Weronika Lajewska, Trond Linjordet

AFFILIATIONS

Information Access & Artificial Intelligence, University of Stavanger, Norway

A **personal knowledge graph** (PKG) is a knowledge graph (KG) where a single individual, called the owner of the PKG, has (1) full read and write access to the KG, and (2) the exclusive right to grant others read and write access to any specified part of the KG. The primary purpose of the PKG is to support the delivery of services that are customized particularly to its owner.

MOTIVATING SCENARIO 1

A **personal trainer assistant**:

- suggests a training plan using personal information
- training recommendations synchronized with the user's calendar
- integrated with external sources of personal information
- integrated with external sources of public data such as Wikidata
- proactive suggestions related to other integrated services in response to user's activity

MOTIVATING SCENARIO 2

A system for **Sharing Health Information**:

- manages and shares patient's sensitive personal information
- gives different health service providers access to pertinent information
- enables different services to stay updated on the current status of the person's disease and treatment
- enables each health service provider to have a pre-defined expected view or a subset of the PKG that is known to be of interest to their technical specialty

DIFFERENT INTERPRETATIONS OF PERSONAL KNOWLEDGE GRAPHS

	PKG (our definition)	PKG (Balog and Kenter, 2019)	Personalized KG
Ownership	Created and maintained by an individual	Created and maintained by an individual	Created and maintained by a service
Public facts	Can incorporate facts from public knowledge graph	Public facts are not explicitly stored, but can be linked	Built with facts from a public/proprietary knowledge graph
Private facts	The owner of the PKG can add private facts (e.g., beliefs) as long as they have the correct format	The owner can add private facts (e.g., beliefs) as long as they are connected to it	Facts that are not of public knowledge cannot be stored (e.g., an individual medication regimen)
Graph structure	Facts do not need to be connected to the user	All facts in the PKG are connected to the user resulting in a spiderweb layout	Facts do not need to be connected to the user

THE PKG ECOSYSTEM

Population:

- **Populating** data from private or public data sources to the PKG.
- **Synchronizing** any modifications made to the extracted private data in the PKG that need to be populated back to the original data source.

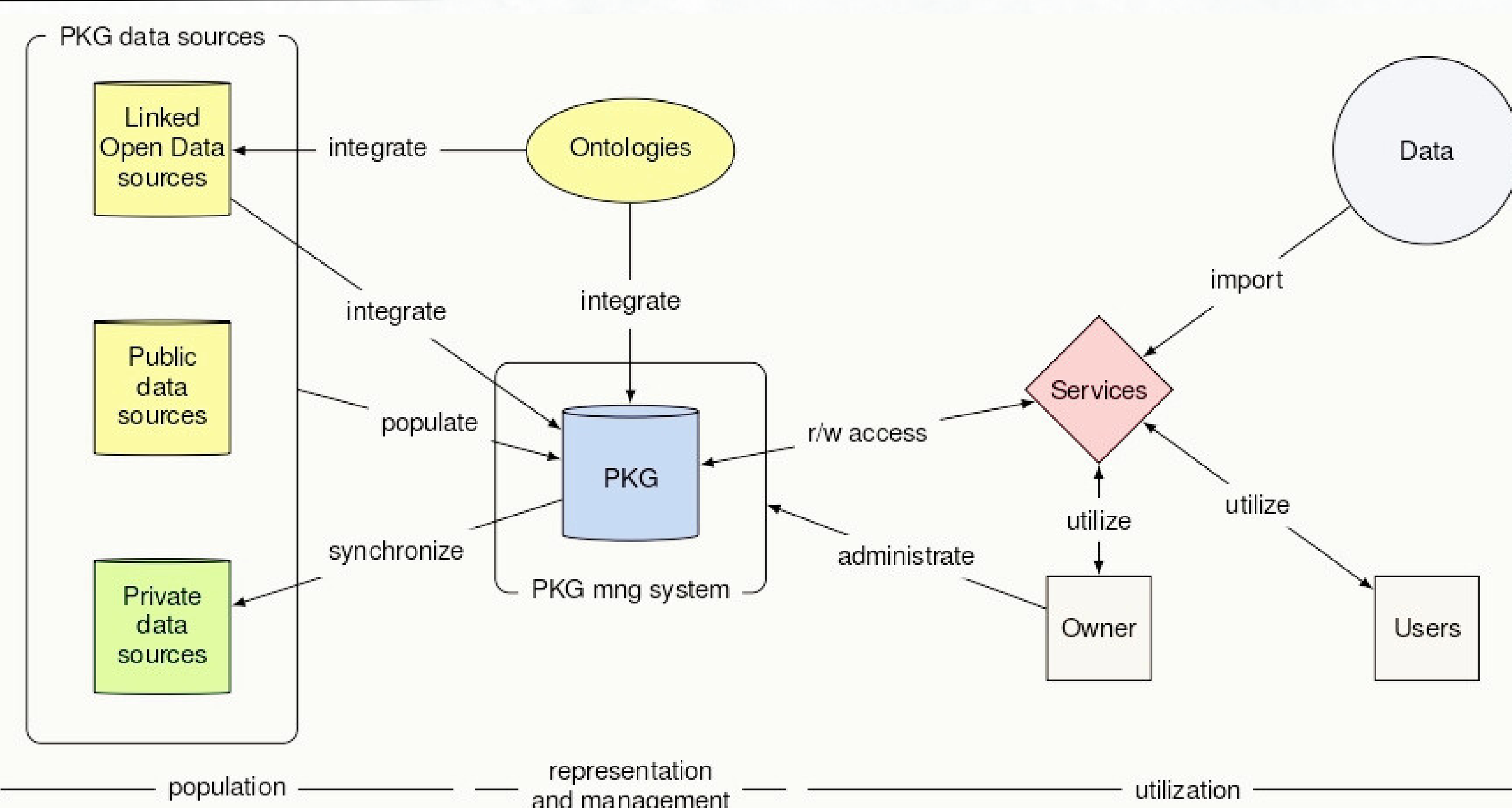
Representation and management:

- Logical **representation**, format, and expressivity of the facts and statements that make out the contents of the PKG.
- **Management**, organization, storage, retrieval, and access control of the content of the PKG.
- **Integrating** PKG with several ontologies to capture the semantics of the facts stored in the PKG.

Utilization:

- Delivering successful personalized services to its owner and users.
- The **administration** concerned with the direct update of data inside the PKG and the control of its access by its owner.

- PAPER DESCRIBING THE PROCESS IN DETAILS
- PAPER BRIEFLY DESCRIBING THE PROCESS



CHALLENGES AND OPPORTUNITIES

- The establishment of PKG standards and their adoption by service providers are the most important open challenges for the realization of a PKG ecosystem.
- New methods are needed for populating a PKG with newly extracted facts while conforming to a pre-determined target ontology and keeping selected external data sources synchronized with the current facts in the PKG.
- New strategies are needed to allow for reasoning over (parts of) the PKG data, both for discovering new facts and for maintenance tasks such as consistency checks.
- The administration process remains a critical challenge for the adoption of PKG especially in the case of sensitive application domains.
- The exploitation of PKG data requires methods for the treatment of missing or inaccessible data.