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.

DIFFERENT INTERPRETATIONS OF PERSONAL KNOWLEDGE GRAPHS

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

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.

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.

