

Figure 2: FuhSen – results screen: 1. Filters or Facets. 2. Category navigation. 3. Result list.

email address, or bank account of a specific organization. It is possible to find such information in internal databases of the law enforcement agencies, on social networks, or even in the Document Web.

Use Case 3: *Where is a certain product being offered?* The objective is to find out in which e-commerce platforms a certain product is being offered at what price.

The web user interface follows a simple search box style where a user can enter search criteria, which could be the name of a person, an organization, or a product. Figure 2 shows how results are displayed to the user grouped by entity type, where the entity summary includes an image and essential properties of the entity. The results are sorted according to the rank value assigned during the search process. Finally, Figure 3 illustrates how facets can be applied for filtering the results. Facets' values are generated automatically from the aggregated results and sorted by frequency. The prototype has been demonstrated to domain experts, where its practical relevance and its general usability have been validated. The platform-as-a-service design opens the possibility to provide *semantic search agents* that automatically search for information, analyze the results, and alert the police authorities when a result is found.

4. CONCLUSIONS

In this paper we demonstrated the foundation for a novel federated, RDF-based hybrid search platform, starting from the design of an architecture, to a comprehensive prototype implementation. The federated, vocabulary-based hybrid search concept constitutes a novel architectural pattern incorporating elements from universal search, semantic integration as well as multi-modal search and retrieval. The aggregation of information from Social Web, Deep Web, Data Web and internal databases represents a novel approach to summarize relevant information about crime investigation.

Although we initially focus on use cases in the criminal investigation domain, we deem that there are numerous further use cases, e.g., related to e-commerce (e.g. price comparison) or social media. Consequently, FuhSen is designed in such a generic, modular and flexible way that it can be adapted easily to other scenarios beyond crime

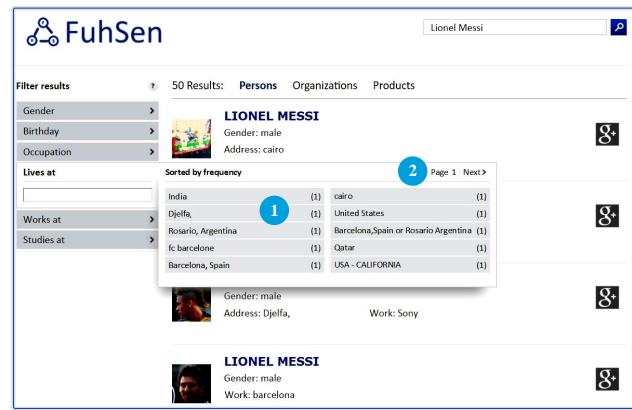


Figure 3: FuhSen – facets selection: 1. Facets values sorted by frequency. 2. Navigation functionality for the facets values.

investigation. Through its clear interface definitions, the flexible vocabulary-based integration models and the modular architecture, new information sources can be plugged-in with minimal effort and the platform can be tailored easily towards new application domains. When applied more widely, this federated search approach can contribute to realizing novel applications and business models, previously prevented by the prohibitive cost of a full data integration.

Acknowledgments. This work was funded by the German Ministry of Education and Research grant no. 13N13627.

References

- [1] J. Iturrioz, I. Azpeitia, and O. Díaz. “YQL as a Platform for Linked-Data Wrapper Development”. In: *Engineering the Web in the Big Data Era*. Springer, 2015.
- [2] V. Lopez et al. “PowerAqua: Supporting users in querying and exploring the semantic web”. In: *Semantic Web 3.3 (2011)*, pp. 249–265.
- [3] G. Montoya et al. “Semlav: Local-as-view mediation for SPARQL queries”. In: *Transactions on Large-Scale Data-and Knowledge-Centered Systems XIII*. Springer, 2014.
- [4] C. Pedrinaci, J. Cardoso, and T. Leidig. “Linked USDL: a vocabulary for web-scale service trading”. In: *The Semantic Web: Trends and Challenges*. Springer, 2014.
- [5] N. Schlaefler et al. “Semantic Extensions of the Ephyra QA System for TREC 2007”. In: *TREC. 2007*.
- [6] S. Shekarpour et al. “Sina: Semantic interpretation of user queries for question answering on interlinked data”. In: *J. of Web Semantics* 30 (2015), pp. 39–51.
- [7] G. Simonini and S. Zhu. “Big data exploration with faceted browsing”. In: *Int. Conf. on High Performance Computing & Simulation (HPCS)*. IEEE, 2015.
- [8] *Social Media Use in Law Enforcement: Crime prevention and investigative activities continue to drive usage*. Tech. rep. LexisNexis Risk Solutions., 2014.
- [9] R. Usbeck et al. “HAWK – Hybrid Question Answering over Linked Data”. In: *ESWC*. Springer, 2015.