ProductKG: A Product Knowledge Graph for User **Assistance in Daily Activities**

Michaela Kümpel^{1,*}, Michael Beetz¹

¹Institute for Artificial Intelligence, University of Bremen, Am Fallturm 1, 28359 Bremen, Germany

Abstract

The Web offers plenty of product information that is valuable for supporting decision processes. Research on Web knowledge acquisition and the Semantic Web has led to the creation of many domain ontologies and Web applications. What still is lacking is a connection of such knowledge to the real world. If object information is linked to environment information, users can get better, more personalised support in their daily activities like shopping or cooking since this enables them to link information about leftover products in the fridge to recipe information or a health profile to products the user is looking at in the store. It has been shown that semantic Digital Twins can successfully link object to environment information that can be used by agents like smartphone or service robot. Such semantic Digital Twins can offer even more services to users if they are connected to product information from the Web.

This work introduces ProductKG, an open-source product knowledge graph integrating modular product information from the Web as well as accurate environment information from a semantic Digital Twin that can be customised for different applications and used devices as an example knowledge graph for assisting users in daily activities. We describe the design process and modularity of the knowledge graph as well as example applications of it, including an Augmented Reality shopping assistant, a dietary recommender and a hands-free recipe application. The modular ontologies enable personalisation of applications as well as accessing object information in relation to the current environment. We evaluate the acceptance of one example application through a user study. ProductKG is publicly available and will be maintained and extended over time in order to facilitate various applications such as in the retail and household domain

Resource Website: https://michaelakuempel.github.io/ProductKG/ Application Website: http://productkg.informatik.uni-bremen.de/ and https://ai.uni-bremen.de/productkg

Keywords

Knowledge Graph, Web Knowledge Acquisition, product ontology, semantic Digital Twin, Knowledge Graph Application

1. Introduction

In the Semantic Web, ontologies and knowledge graphs have been a research focus for a while, resulting in partially interlinked domain ontologies with different consumer applications. It has been shown that knowledge graphs are a powerful tool that can be used to efficiently answer questions. However, there is a need for research on linking knowledge graphs to real world

D 0000-0002-0408-3953 (M. Kümpel); 0000-0002-7888-7444 (M. Beetz)



^{© 2023} Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0). CEUR Workshop Proceedings (CEUR-WS.org)

Ontology Showcase and Demonstrations Track, 9th Joint Ontology Workshops (JOWO 2023), co-located with FOIS 2023, 19-20 July, 2023, Sherbrooke, Québec, Canada.

^{*}Corresponding author.

michaela.kuempel@uni-bremen.de (M. Kümpel); beetz@cs.uni-bremen.de (M. Beetz)

systems. If knowledge graphs can be grounded in the real world, the contained knowledge becomes perceivable and actionable to users. While we believe that this connection of object to environment information is beneficial for many different domains, in this work we concentrate on product information, which is due to

- Environment structure. Products often hold a barcode as identifier and can usually be found in very structured environments: they are stored in shelves or cabinets, which helps in identifying perceived objects.
- Availability of environment information. Digital agents can already create environment models of retail stores and household environments autonomously [1, 2].
- Availability of Web Information. Product information as well as product images are available in many web stores and consumer websites.

It has been shown that *semantic Digital Twins* (semDTs) can successfully be used for linking object to environment information in the retail domain for customer support [3]. The concept of a *Digital Twin*, "the digital equivalent to a physical product" was introduced by Grieves in 2003 [4]. Using Digital Twins to create exact virtual representations of the real world has been a technology trend in digitisation of manufacturing and industry [5]. A *semantic Digital Twin* can be described as the semantic connection of environment information into a digital representation of an environment. While a semDT offers detailed environment information, it only offers limited object or product information from connected enterprise management systems. We therefore extend the idea and link semDTs to a *knowledge graph* of product information acquired from Web sources. In this work, we refer to a *knowledge graph* as an aggregation of ontologies comprising Web information that allow reasoning about the contained facts. With this connection we get access to both detailed environment and vast product information.

This work introduces ProductKG, a knowledge graph that consists of semantic product information from the Semantic Web and the World Wide Web in modular ontologies that are linked to environment information which can be combined and customised for various applications. It enables a customer browsing the aisles with their smartphone to highlight all deodorants that were dermatologically tested but also visualise nutritional information for a product in the fridge, for example. In this work we describe the creation of the ProductKG knowledge graph and its ontologies. We showcase the benefits of using ProductKG in example applications in



Figure 1: Example applications enabled by ProductKG: left: smart glass shopping support, middle: smart glass recipe assistant, right: smartphone recipe and shopping support.

retail and household environments as depicted exemplary in Figure 1. The applications use the same basic user profile with different implementations like highlighting product preferences, recommending recipes for available products or showing preferred information such as product labels. The novelty of ProductKG lies in the connection of exact location information from a semantic Digital Twin to modular product information acquired from the Web into a knwoldge graph, leading to various applications with consumer focus as an embodiment of the contained knowledge.

The main contributions of this paper are: first, we create ProductKG with its modular ontologies based on Web information, link it to other Semantic Web sources and make it publicly available. Second, we integrate semDT environment information in such a way that it can be used by AR devices as well as robots. Third, we showcase the applicability of ProductKG in daily environments on smartphone and smart glasses, thus making ProductKG perceivable and actionable. These contributions are validated by a system usability study for one of the applications.

2. Related Work

On the one hand, recent research has focused on standardised mapping, in particular of retail environments. The benefits of using standardised maps has been identified in [6]. The idea of creating semantic Digital Twins of retail environments as standardised maps has been proposed in [3] for shopping assistance or shop-floor assistance for product refilling. It has additionally been shown how semDTs can be used in retail applications to route a customer to a searched product on different devices such as smartphone and robot [2]. While the shown applications reveal the power of semantic Digital Twins and the included environment information, they lack detailed product information and cannot easily be extended to other applications like finding substitute products in a store or recipes one can prepare with a given product.

On the other hand, there has been research on creation of product knowledge graphs (e.g. [7, 8]) and food product knowledge graphs, resulting in partially interlinked domain ontologies with consumer applications ranging from agricultural over health specific to recipe centered (e.g. [9, 10, 11]). Although the created ontologies, knowledge graphs and their exemplary applications are impressive, they do not connect the contained object or product information to environment information.

This work aims at bringing together both research topics by interlinking semantic Digital Twins with a product knowledge graph.

3. Acquired Knowledge

Knowledge graphs can be created a) manually (which will not be further discussed in the following), b) by accessing structured external (Web) sources and c) by accessing unstructured external (Web) sources [12]. ProductKG consists of a location ontology created by accessing a robot map (as described in [1]) as well as modular product information ontologies that contain structured or unstructured Web information as described in the following.

3.1. Structured External Sources

In order to generalise applications for the use on different devices and in differing applications, we need to have precise, interchangeable environment information. A standardized map is also beneficial for localization of products in a store. We create a location ontology that is linked to precise environment information in a semantic Digital Twin environment model. A location ontology can be generated for any indoor environment following the approach described in [1] and [3]. We additionally derive product dimension information from the perception system. The robot uses its RGB and RGB-D sensors to identify the amount of consecutive products and calculate product width, height and depth.

The positions in the semantic Digital Twin use a fixed reference frame with a given origin, usually set to a randomly chosen corner of the room. In contrast to this, Augmented Reality (AR) devices use the varying device position when starting an application as origin and display its digital content relative to the device position. To solve this discrepancy for a use of the location ontology in both robot and AR applications, we encode all product locations in the location ontology relative to shelf, table or other object positions as described in more detail in [2]. In the AR shopping applications we then use spatial perception and match the world origin of the game to the reference frame origin of the semantic Digital Twin. We also use world anchors for locking positions in the physical space, a technique that has proven effective in other work [13]. World anchors need to be created once and will be loaded in every subsequent run of the application. Thus, product positions can be inferred relative to world anchors and digital content can be displayed.

3.2. Structured Web Information

ProductKG integrates and links to structured Web information. On the one hand, we integrate parts of structured Web information like the FoodKG ontologies [11] for a product-specific recipe recommendation. This is done by reusing the food product classification of the FoodOn ontology [9], which we intertwine with the product classification in the product taxonomy. ProductKG also integrates nutrition information from the Compositional Dietary Nutrition (cdno) ontology [14] as well as nutritional product information like Nutri-Score, product labels or packaging size for food products from Open Food Facts ¹. For food classes, we access product class nutrition information from the FoodData central ² from the U.S. Department of agriculture. The information is stored in a food-nutrition ontology, which then links to the FoodOn classification and therefore can be used for food products in the product taxonomy.

On the other hand, we interlink ProductKG to existing data sources by re-using many existing URIs and standard vocabulary as proposed in the Linked Data standards by Bizer, Heath et al. [15]. We do not, however, use the owl:sameAs property but the oboInOwl:hasDbXref annotation property to interlink two data sources, as proposed by the gene ontology (GO) consortium [16]. This is due to the fact that owl:sameAs increases file sizes since it creates duplicate entries while the oboInOwl:hasDbXref not only avoids duplicate entries but also looks more descriptive to a user browsing the ontologies using Protégé. Both the GO and

¹Open Food Facts: https://world.openfoodfacts.org/

²Food Data central: https://fdc.nal.usda.gov/

the FoodOn workgroup are part of the OBO foundry [17], a community focussing on the development of interoperable ontologies for the biological sciences. We chose to reuse many of the OBO foundry ontologies (besides FoodOn and cdno we link to the human disease ontology (doid) [18, 19] and the human phenotype ontology [20]) since they are openly accessible, are already following the same upper ontologies and have a strong community.

The product classification is further linked to Wikidata and the Product Types Ontology³ wherever possible. We also use the Good Relations web vocabulary for E-Commerce⁴. Furthermore, ProductKG links to chemical information about allergens in the Chemical entities of biological interest (ChEBI)⁵ ontology. The allergen ontology additionally links to other Semantic Web sources like the National Drug File (NDRF) of BioPortal⁶ or Drugbank⁷. The advantage of using many Linked Data sources lies in its adaptability to changes. If a concept is changed in a linked source, the new information is immediately available in applications.

3.3. Unstructured Web Information

For creation of the ProductKG ontologies we additionally acquire unstructured Web information to create structured knowledge sources that machines can query. In order for ProductKG to be usable in various environments, it needs a general product taxonomy that can be used for different domains of daily activities like at home, in drugstores and grocery stores. Most techniques for taxonomy creation aim at properly modeling a domain by using a pattern-based approach for taxonomy induction [21] to learn simple but precise contextual patterns in a Web document. For a more generalised approach we use Information Extraction [22] techniques on online store sitemaps of different retail domains to automatically create a more general product taxonomy.

Since online stores usually offer a sitemap in their robots file⁸, we use this sitemap for Information Extraction to avoid spamming of websites and be more time-efficient. Sitemaps of online stores are usually hidden .xml files that consist of a list of Uniform Resource Locators (URLs) of children websites so that crawlers can easily access them. For a general product classification we create a product class structure from sitemaps of two online stores representing different retail sectors, namely Aldi-Nord (grocery store) and dm (drugstore). We then use Owlready to generate ontologies out of the extracted information. The resulting product taxonomies are merged into one global product taxonomy that additionally integrates parts of the FoodOn food product taxonomy. For this, extracted product categories are standardised to lower case words while special characters and numbers are deleted. This allows for an accurate matching of categories of different stores as will further be explained in Section 4.3.

We further use Information Extraction to acquire product ingredient information from online stores for use in the ingredient ontology and information like brand, awarded labels, weight,

³Product Types Ontology: http://www.productontology.org/

⁴Good Realtions vocabulary: http://www.heppnetz.de/ontologies/goodrelations/v1

⁵Chemical entities of biological interest: https://www.ebi.ac.uk/chebi/

⁶NDRF at BioPortal: https://bioportal.bioontology.org/ontologies/NDFRT/

⁷Drugbank: https://go.drugbank.com/

⁸A robots file is the /robots.txt file located at the root of a website with instructions for web robots on how to crawl the website and which information can be axtracted.

filling capacity and country of origin from Codecheck⁹, a consumer-oriented product information website. Again, we standardise the extracted information as before and manually add additional semantic information such as linking a specific label to a label category. What is more, product ingredients are classified as allergens according to information from the U.S. Food & Drug administration (FDA) website¹⁰.

4. Composition of ProductKG

We continue to explain the contained ontologies in ProductKG as well their accordance to the FAIR (Findable, Accessible, Interoperable, Reusable) principles [23] and how ProductKG can be extended in the following.

4.1. Ontology Alignment

As depicted in Figure 2, the product builds the core of ProductKG. The modular ontologies in ProductKG have different foci that can be used to retrieve additional product information for different applications. All ontologies contained in ProductKG with their most relevant properties and their current use in applications as well as the number of axioms are detailed in Table 1.

¹⁰FDA website: https://www.fda.gov/



Figure 2: Overview of ProductKG ontologies and their links to external sources.

⁹Codecheck website: https://www.codecheck.info/

Table 1

Ontologies contained in ProductKG with their most relevant properties and their use in applications. *Italic* information sources belong to ProductKG while all other mentioned sources are external.

Ontology	Information Sources	Offered Properties	Axioms	Apps
Product Taxonomy	Good Relations vocabulary wikidata Aldi Nord, dm Food Ontology semDT	gr:hasEAN_UCC-13 wikientry trust:source oboInOwl:hasDbXref pathToCadModel	8,603	all
Location Ontology	semDT	loc:has_stock loc:has_price loc:stored_in/ stored_on	7,721	App 1 routing [2]
Dimension Ontology	QUDT Good Relations vocabulary	qudt:unit qudt:hasQuantity gr:depth/ height/ width/ weight	4,312	App 2 shelf refill [24]
Brand Ontology	Good Relations vocabulary	gr:hasBrand gr:name has_StoreBrand has_subbrand	5,669	Арр 1 Арр 2
Packaging Ontology		has_PackagingMaterial has_RecyclingProperty	5,435	Арр 2
Label Ontology	Friend of a Friend	foaf:depiction has_label	3,341	Арр 2
Nutrition Ontology	QUDT	qudt:has_quantity has_nutrient	95,192	Арр 3
Symptom Ontology	Disease Nutrient	has_symptom possible_treatment	1,487	Арр 3
Disease Ontology	Ingredient Symptom	triggers symptom_of	582	Арр 1 Арр 3
Ingredient Ontology		has_ingredient has_trace	18,050	Арр 1 Арр 3
Recipe Ontology	FoodKG QUDT	food:isRecommendedForCourse hasMealType hasIngredient qudt:has_Quantity	295,231	App 3 Recipe App
Substitute Ontology	Ingredient QUDT	has_substitute has_purpose qudt:has_Quantity	2,0580	App 3 Recipe App
Allergen Ontology	Disease FOAF Ingredient	has_trigger foaf:has_depiction owl:same_as	1,327	Арр 1 Арр 3

One major challenge in ontology alignment of ontologies with such different foci is the use of diverse upper ontologies. While ProductKG is based on the DOLCE+DnS Ultralite (DUL) ontology with its descriptions of relations of actions, objects and agents, the OBO foundry ontologies are based on the Basic Formal Ontology (bfo) defining occurrents and continuants. Thus, external concepts like processes that are defined as occurents in bfo need to be aligned with the definitions of tasks, actions and processes as events from DUL. To simplify ontology reuse in ProductKG, we focus on the domain and solely integrate product information. In particular, the main DUL class ProductKG is based upon is *PhysicalObject* - the class of all objects perceivable in an environment. In DUL, a *PhysicalObject* \in *Object* \in *Entity*. In bfo, a *FoodProduct* \in *FoodMaterial* \in *MaterialEntity* \in *Continuant* \in *Enity*. In ProductKG we align these by stating that

 $FoodProduct(bfo) \in groceries \in ProductOrService(gr) \in PhysicalObject(DUL)$ a food product of the Food Ontology is a subclass of groceries in ProductKG, which we define as a Product or Service of the Good Relations vocabulary, which then belongs to the physical object class of DUL.

4.2. Accordance to FAIR principles

The FAIR guiding principles for scientific data management and stewardship [23] were proposed in 2016 in order to improve the reusability of data holdings. They introduce four principles: Findability, Accessibility, Interoperability, and Reusability that we apply to ProductKG as explained in the following:

- Findability: ProductKG and its application page are openly accessible and can be found at https://michaelakuempel.github.io/ProductKG/. We use persistent URLs (http://purl. org/ProductKG/) for the contained ontologies that redirect the user to the current storage location of the ontologies.
- Accessibility: ProductKG is publicly available with all ontologies, a link to its triple store SPARQL endpoint¹¹ as well as an grlc[25] RESTful API for the example queries¹² demonstrated in the shopping assistant applications.
- Interoperability: Section 3.2 describes the different vocabularies that are used or linked to in ProductKG. Hence, applications built on those ontologies can easily integrate ProductKG information.
- Reusability: Entities in the ProductKG ontologies use annotation properties like language labels (which currently are implemented in english and german), comments and links to term defining webpages such as wikidata or the product ontology as descriptors. All ProductKG ontologies come with a clearly stated licence, can be downloaded and reused.

4.3. Extending ProductKG

ProductKG can easily be extended if new sources are linked to product names. Using product names for classification has been shown to be quite efficient for linking. For ontologies,

¹¹The ProductKG API can be accessed at this SPARQL endpoint: https://api.krr.triply.cc/datasets/mkumpel/ ProductKG/services/ProductKG/sparql

¹²Example queries can be accessed online: http://grlc.io/api/michaelakuempel/ProductKG/SPARQLfiles/

Word2Vec can be used to expand the contained knowledge by learning sibling classes [26] and for linking taxonomies one could use either Word2Vec or Node2Vec with good results. If we want to link quite different ontologies such as taxonomies to a brand ontology, the domain of interest changes from linking two "shampoo" classes to linking one "shampoo" class to "Elvital Dream Length Super Strengthening Shampoo", for example. Here, the distinct product name consists of multiple words. Doc2Vec has been shown to efficiently handle multiple words with context [27]. Therefore, we use Doc2Vec for linking of the different ontologies to the product taxonomy.

Even product master data can easily be linked to ProductKG, which is important if different users want to access ProductKG information. For this, we propose to do a preprocessing to have all product names in lower case format. We tested linking product master data containing about 300,000 product names and received a top-1 accuracy of 80% and a top 5 accuracy of 97%. What is more, if the algorithm suggested to have an accuracy of more than 90%, it was actually true in 95% of tested cases.

5. Consumer Applications of ProductKG

Using ProductKG, we can answer many complex questions about products, their ingredients, contained allergens, nutrition values, labels, brands or recipes that can be prepared with them, for example. The connection to the semantic Digital Twin additionally enables to locate products in the current environment. It has already been shown in previous work how semantic Digital Twins can successfully be used for shopping assistants to help customers find searched products [2]. The triple store used for storage of the knowledge graph offers a SPARQL endpoint for information access that can be accessed by different agents like robot or smartphone.

The following applications demonstrate the applicability of the ProductKG knowledge graph as one source of data for a range of consumer applications. In omnichannel applications with one source of data, it does not matter if the consumer accesses the information through a Web interface, a smartphone, smart glasses or via interaction with a robot. The following applications show how the same data can be accessed through various channels for different applications. The used devices in the applications are exchangeable.

5.1. Application 1: Preference Visualisation in AR Using HoloLens

This openly accessible application¹³ highlights all objects that do not comply with a set preference on a HoloLens as depicted in Figure 1 on the left. A user can specify a preference such as an unwanted ingredient and all products that contain the ingredient get a digital overlay, a red X.

To localize the HoloLens in the retail laboratory we use spatial perception and match the world origin of the Unity game to the reference frame origin of the robot environment map as discussed in 3.1. To load an environment map, a user has to scan a QR code that loads the semantic Digital Twin map including the parent anchor. Product positions are inferred relative to this parent anchor. Each product in the game is created as a X game object named according

¹³HoloLens application: https://github.com/michaelakuempel/HoloPreferenceDemo

to the ProductKG product names. Thereby, objects can be set visible or invisible automatically according to the query result.

A consumer in the store will see a menu on the wall, displaying options of preferences like "vegan", "eco-friendly" or "no alcohol". Once a preference is chosen, all opposing products will be highlighted as depicted in Figure 1.

5.2. Application 2: Information Visualisation in AR Using a Smartphone.

This application highlights interesting product information using a smartphone. It uses image targets for image recognition. Therefore, we rely on images or barcodes of products available in daily environments. Hence this application works more reliable in the more static retail environment than in dynamic household environments.

A consumer using the app can scan a product they are interested in to instantly see interesting product information like which label the recognised product has as well as an information slider displaying the name of the recognized product. By clicking on or sliding of the product name, additional product information like contained ingredients and allergens or recipes that can be prepared with the product can be listed. Aside from that, the product can be added to the shopping cart. The smartphone application to visualize interesting product information is depicted in Figure 1 on the right for a recipe recommender and a shopping assistant.



Figure 3: Personalised nutrition recommender.

5.3. Application 3: Personalised Nutrition Recommender

ProductKG offers a basic user profile ontology that can be integrated and instantiated in ontologies like the food-nutrition or ingredient ontology for personalised applications. Figure 3 shows an example nutrition recommender application based on personal dietary preferences¹⁴. A user can (but does not need to) enter personal information like weight, age, gender and daily activity level. The data is not saved, but the user is instantiated based on the given input values. In Figure 3, a user is searching for products that are rich in vitamin D, nutritional values of coffee and nutrients and products that might treat a headache. The results pages show products in accordance to the preferences and query. Based on this, the user can proceed to see detailed product information or what recipes could be prepared with it. Through connected semDTs, users can also see where a searched product is available or compare prices.

6. System Usability Evaluation

For an assessment of system usability we performed a user study based on the system usability scale, which we extended for a question about the helpfulness of the application, i.e. if the users think that the dietary recommendation will help them in achieving their goals. We performed the user study with n=24 participants by promoting the website and randomly asking website visitors for their feedback based on the following hypotheses:

- We expect that the application is easy to use.
- We expect that the idea of the website is understood quickly.
- We expect that the functionality of the website is understood easily.
- We expect the overall application and its subpages to be perceived equally complex.

The performed user study yielded some interesting results, some of which are visualised in Figure 4. Overall, the user study results suggest that users would reuse the application.



¹⁴The Web Applications are available at: http://productkg.informatik.uni-bremen.de/

Figure 4: Evaluation results for examining App usability.

- As expected, the users did find the application in general easy to use. Surprisingly, this is only true for the general application. Some subpages were assessed as being too complex.
- The users were able to quickly grasp the application and its subpages. Again, the overall application received better results than its subpages.
- The functionality of the app was evaluated as easily understandable both for the general application and its subpages.
- It is interesting to see that the results of the application in general and the subpages offering different information differ. Users were very interested in the symptom page and used it more than other pages.
- The nutrition recommender subpage was evaluated as being very helpful.
- Most of the users would reuse the app. The symptom subpage received most interest. Users said that they were most likely to reuse this part of the application. Those users that would not reuse the app mainly were disappointed in the design of the application.

7. Results and Future Work

This work describes the ProductKG knowledge graph with all its ontologies and points out the benefits of linking a knowledge graph to a semantic Digital Twin, making it perceivable and actionable. All sources used for information acquisition and creation of the contained ontologies are described. We show the applicability of the knowledge graph in various applications. Having a publicly available knowledge graph makes it accessible to different machines, which we demonstrate by querying ProductKG by smart glass (HoloLens) and smartphone. The example applications emphasise the value of the connection of semantic product information to location information, while the power of the applications stem from the use of modular ontologies that can be personalized and queried through a single access point. ProductKG offers:

- A product taxonomy that can be reused for applications in different daily environments.
- A framework of modular ontologies that can be accessed for different consumer needs and applications.
- A connection to existing ontologies like wikidata, FoodOn or ChEBI, allowing for further applications based on the contained knowledge.
- A connection to exchangeable, standardised Digital Twin environment models, thereby enabling environment-dependent applications, bringing the knowledge to the user.

We believe that ProductKG is a powerful example of connecting knowledge graphs to semantic Digital Twins and will continue to use this approach for different domains such as bookstores as well as more user applications in daily environments. Future work will additionally focus on an integration of the different applications into a single daily activity application. Additionally, we hope to add more semDTs that are openly accessible and we plan on using knowledge graphs for enabling higher robot autonomy in daily environments.

Acknowledgments

We thank Naser Azizi and Sorin Arion for their support towards this work.

The research reported in this paper has been partially supported by the German Federal Ministy of Education and Research; Project-ID 16DHBKI047 "IntEL4CoRo - Integrated Learning Environment for Cognitive Robotics", University of Bremen, the Federal Ministry for Economic Affairs and Energy BMWi within the Knowledge4Retail project, subproject semantic Digital Twin 01MK20001M (https://knowledge4retail.org) as well as the German Research Foundation DFG, as part of Collaborative Research Center (SFB) 1320 "EASE - Everyday Activity Science and Engineering", University of Bremen (http://www.ease-crc.org/). The research was conducted in subproject R04 "Cognition-enabled execution of everyday actions".

References

- M. Beetz, S. Stelter, D. Beßler, K. Dhanabalachandran, M. Neumann, P. Mania, A. Haidu, Robots Collecting Data: Modelling Stores, Springer International Publishing, Cham, 2022, pp. 41–64. doi:10.1007/978-3-031-06078-6_2.
- [2] M. Kümpel, J. Dech, A. Hawkin, M. Beetz, Robotic shopping assistance for everyone: Dynamic query generation on a semantic digital twin as a basis for autonomous shopping assistance, in: Proceedings of the 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2023), London, United Kingdom, 2023, pp. 2523–2525.
- [3] M. Kümpel, C. A. Mueller, M. Beetz, Semantic digital twins for retail logistics, in: M. Freitag, H. Kotzab, N. Megow (Eds.), Dynamics in Logistics: Twenty-Five Years of Interdisciplinary Logistics Research in Bremen, Germany, Springer International Publishing, Cham, 2021, pp. 129–153. doi:10.1007/978-3-030-88662-2_7.
- [4] M. Grieves, Virtually perfect: driving innovative and lean products through product lifecycle management, Space Coast Press, 2011.
- [5] P. Augustine, The industry use cases for the digital twin idea, in: Advances in Computers, volume 117, Elsevier, 2020, pp. 79–105.
- [6] Z. Davis, M. Hu, S. Prasad, M. Schuricht, P. Melliar-Smith, L. E. Moser, A personal handheld multi-modal shopping assistant, in: International conference on Networking and Services (ICNS'06), IEEE, 2006, pp. 117–117.
- [7] T. Lee, I.-h. Lee, S. Lee, S.-g. Lee, D. Kim, J. Chun, H. Lee, J. Shim, Building an operational product ontology system, Electronic Commerce Research and Applications 5 (2006) 16–28.
- [8] N. Zalmout, C. Zhang, X. Li, Y. Liang, X. L. Dong, All you need to know to build a product knowledge graph, in: Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining, 2021, pp. 4090–4091.
- [9] D. M. Dooley, E. J. Griffiths, G. S. Gosal, P. L. Buttigieg, R. Hoehndorf, M. C. Lange, L. M. Schriml, F. S. Brinkman, W. W. Hsiao, Foodon: a harmonized food ontology to increase global food traceability, quality control and data integration, npj Science of Food 2 (2018) 1–10.
- [10] J. Cantais, D. Dominguez, V. Gigante, L. Laera, V. Tamma, An example of food ontology for diabetes control, in: Proceedings of the International Semantic Web Conference 2005 workshop on Ontology Patterns for the Semantic Web, 2005, pp. 1–9.
- [11] S. Haussmann, O. Seneviratne, Y. Chen, Y. Ne'eman, J. Codella, C.-H. Chen, D. L. McGuinness, M. J. Zaki, Foodkg: a semantics-driven knowledge graph for food recommendation,

in: International Semantic Web Conference, Springer, 2019, pp. 146-162.

- [12] N. Heist, S. Hertling, D. Ringler, H. Paulheim, Knowledge graphs on the web–an overview, Knowledge Graphs for eXplainable Artificial Intelligence: Foundations, Applications and Challenges (2020) 3–22.
- [13] A. Jakl, L. Schöffer, M. Husinsky, M. Wagner, Augmented reality for industry 4.0: Architecture and user experience., in: FMT, 2018, pp. 38–42.
- [14] L. Andrés-Hernández, A. Baten, R. Azman Halimi, R. Walls, G. J. King, Knowledge representation and data sharing to unlock crop variation for nutritional food security, Crop Science 60 (2020) 516–529.
- [15] C. Bizer, R. Cyganiak, T. Heath, et al., How to publish linked data on the web (2007).
- [16] G. O. Consortium, The gene ontology resource: 20 years and still going strong, Nucleic acids research 47 (2019) D330–D338.
- [17] B. Smith, M. Ashburner, C. Rosse, J. Bard, W. Bug, W. Ceusters, L. J. Goldberg, K. Eilbeck, A. Ireland, C. J. Mungall, et al., The obo foundry: coordinated evolution of ontologies to support biomedical data integration, Nature biotechnology 25 (2007) 1251–1255.
- [18] L. M. Schriml, C. Arze, S. Nadendla, Y.-W. W. Chang, M. Mazaitis, V. Felix, G. Feng, W. A. Kibbe, Disease ontology: a backbone for disease semantic integration, Nucleic acids research 40 (2012) D940–D946.
- [19] L. M. Schriml, J. B. Munro, M. Schor, D. Olley, C. McCracken, V. Felix, J. A. Baron, R. Jackson, S. M. Bello, C. Bearer, et al., The human disease ontology 2022 update, Nucleic acids research 50 (2022) D1255–D1261.
- [20] P. N. Robinson, S. Mundlos, The human phenotype ontology, Clinical genetics 77 (2010) 525–534.
- [21] H. Davulcu, S. Koduri, S. Nagarajan, Datarover: a taxonomy based crawler for automated data extraction from data-intensive websites, in: Proceedings of the 5th ACM international workshop on Web information and data management, 2003, pp. 9–14.
- [22] O. Etzioni, M. Banko, S. Soderland, D. S. Weld, Open information extraction from the web, Commun. ACM 51 (2008) 68–74. doi:10.1145/1409360.1409378.
- [23] M. D. Wilkinson, M. Dumontier, I. J. Aalbersberg, G. Appleton, M. Axton, A. Baak, N. Blomberg, J.-W. Boiten, L. B. da Silva Santos, P. E. Bourne, et al., The fair guiding principles for scientific data management and stewardship, Scientific data 3 (2016) 1–9.
- [24] A. Cavallo, M. Costanzo, G. De Maria, C. Natale, S. Pirozzi, S. Stelter, G. Kazhoyan, S. Koralewski, M. Beetz, Robotic Clerks: Autonomous Shelf Refilling, Springer International Publishing, Cham, 2022, pp. 137–170. doi:10.1007/978-3-031-06078-6_6.
- [25] A. Meroño-Peñuela, R. Hoekstra, grlc Makes GitHub Taste Like Linked Data APIs, in: The Semantic Web: ESWC 2016 Satellite Events, Heraklion, Crete, Greece, May 29 – June 2, 2016, Springer, 2016, pp. 342–353. doi:10.1007/978-3-319-47602-5_48.
- [26] G. Wohlgenannt, F. Minic, Using word2vec to build a simple ontology learning system., in: ISWC (Posters & Demos), 2016.
- [27] J. H. Lau, T. Baldwin, An empirical evaluation of doc2vec with practical insights into document embedding generation (2016).