NonFoodKG: A Non-Food to Food Product Knowledge Graph for Consumer Applications in Retail Environments

Michaela Kümpel [a,*] and Michael Beetz [a]

[a] *Institute for Artificial Intelligence, University of Bremen, Bremen, Germany*
E-mail: [michaela.kuempel,beetz]@cs.uni-bremen.de

Abstract. The Web offers plenty of product information that is valuable for supporting customer decision processes. On the one hand, food product knowledge has been shown to be useful for consumer applications like recipe or dietary recommendation, for example. On the other hand, product knowledge can be extremely useful for applications in retail stores when it is linked to a semantic environment map. That way, it can be used to help customers find products they are searching for or highlight interesting product information like awarded labels or contained ingredients and allergens. Unfortunately, the Semantic Web lacks a semantically enhanced knowledge graph unifying non-food product information while the existing shopping applications lack a standardised connection to environment information.

This work proposes NonFoodKG, an open-source non-food to food product knowledge graph integrating modular product information from the Web as well as accurate environment information from sensor data that can be customised for different applications and used devices. We describe the design process and modularity of the knowledge graph as well as example applications of it, including an Augmented Reality shopping assistant to highlight all products containing a given preference like an ingredient/ allergen or label and a dietary recommendation. We show how a routing application to find product destinations can easily be used by different agents like smartphone or robot. Thus, NonFoodKG can be seen as an example of one point of data access in digitisation of retail stores towards omnichannel applications. NonFoodKG links to environment information, which we create using the KnowRob knowledge processing framework. This connection enables agents to link product to action knowledge and thereby generalise the product information to be used in different domains while the modular ontologies enable personalisation of applications. NonFoodKG 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://k4r-iai.github.io/NonFoodKG/

Keywords: Knowledge Graph, Web Knowledge Acquisition, non-food ontology, semantic environment map, Retail application, Robot Application, Augmented Reality application

1. Introduction

In the last decades we have seen many interesting consumer applications in the offline retail domain. The applications range from mobile [1, 2] or robot shopping assistance [3, 4] over robotic depalletizing and shelf replenishment [5, 6] to click-and-collect services and smart shopping carts.

*Corresponding author. E-mail: [michaela.kuempel,beetz]@cs.uni-bremen.de.

1570-0844/$35.00 © 0 – IOS Press and the authors. All rights reserved

Problem While these applications are impressive and a first step towards increased customer decision support, they often are too customised, meeting only certain consumer profiles. What is more, the same consumer will demand different product information like its availability and price, ingredients that might cause allergic reactions or harm the environment, recommendations for complementary products as well as convenient return policies, depending on the temporal relation to the actual purchase (before, during or after) [7] and the users' personal relationship to the device used for information retrieval [8]. In order for applications to turn into omnichannel applications providing a seamless shopping experience that is independent of the used device or environment [9], we suggest that they should: a) share one source of data, b) be generalised and hence can be instantiated for a personalisation that takes into account the temporal relation to a purchase and c) include semantically encoded environment information.

Proposed Approach While online stores often include Web information to further support customer decision processes and make recommendations based on them [10], offline retail stores lack this additional knowledge source. Hence, offline retailers apparently use different data sources than their online counterparts. The latest developments of *the Semantic Web* suggest to use ontologies as such knowledge sources. *The Semantic Web* [11] is an extension of the current Web [12], aiming at structuring content of Web pages by using standardised formats to represent entities, their properties and relations in ontologies that can be understood by robotic or software agents. In the Semantic Web, food ontologies have been a research focus for a while, resulting in partially interlinked domain ontologies with consumer applications ranging from agricultural over health specific to recipe centered (e.g. [13–15]). Contrary to this, non-food product information on the Semantic Web has only been mentioned as support for food ontologies so far [16]. We claim that non-food product ontologies can lead to as many successful applications as food ontologies since consumers are just as interested in allergens in cosmetic products or detergent as they are in food intolerances, for example. The Semantic Web lacks non-food ontologies as well as an information unifying non-food *knowledge graph* that could be used as a single source of data for such applications. In this work, we refer to a *knowledge graph* as an aggregation of ontologies comprising Web information that allow reasoning about the contained facts. A non-food knowledge graph can enable a customer browsing the aisles with their smartphone to highlight all deodorants that were dermatologically tested and come in environmentally friendly packaging, for example. Additionally, a knowledge graph offers a single access point for applications and therefore solves the problem of distributed data mentioned above in a).

This work introduces NonFoodKG, a knowledge graph that consists of semantic non-food and food product information from the Semantic Web and the World Wide Web in modular ontologies that can be combined and customised for various applications. To further meet consumer needs in temporal relation to the purchase as stated in b), we additionally make NonFoodKG accessible through *KnowRob* [17, 18]. *KnowRob* is a knowledge processing framework designed for cognition-enabled robots that also integrates the socio-physical model of activities (SOMA) [19], enabling robotic or software agents to reason about environments and actions, amongst others. Thereby, the agent can assume the temporal relation to purchase (e.g. only during a shopping activity the agent/ device will be in a retail store) or the role of an object or a person at a given time.

We additionally connect NonFoodKG to a *Digital Twin* store model containing semantic information about objects a robotic agent perceives in a retail environment, thereby creating a *semantic Digital Twin* [20]. The concept of a *Digital Twin*, *"the digital equivalent to a physical product"* was introduced by Grieves in 2003 [21]. Using

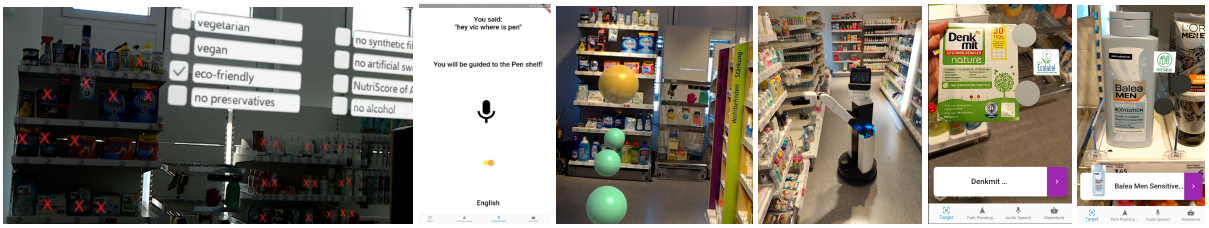

Fig. 1.: Example shopping applications enabled by NonFoodKG:
left: smart glass shopping support, middle: routing application on HoloLens and robot, right: smartphone shopping support.

Digital Twins to create exact virtual representations of the real world has been a technology trend in digitization of manufacturing and industry [22]. We refer to a *semantic Digital Twin* as the semantic connection of environment information into a digital representation of a retail store including offered products, their properties and relations that can be perceived with a sensor. This unique connection of product to environment information that can be understood by machines addresses the problem in c).

Contribution In this work we describe the creation of the NonFoodKG knowledge graph and its ontologies. We showcase the benefits of using a semantic Digital Twin knowledge base that connects location to product information in example applications as depicted in Figure 1. The applications use the same basic user profile with different implementations like highlighting preferences (Fig. 1, left) or showing preferred information (Fig. 1, right), referencing the need to adapt to different consumer needs. We also show how the same information can be used in similar applications to route a customer to a searched product on different devices (smartphone, smart glasses and robot, Fig. 1, middle and Fig. 4), referencing differing personal relationships of consumer and device.

The novelty of NonFoodKG lies in its connection of exact location information to modular product information from the Web, leading to various applications with consumer focus as an embodiment of the contained knowledge, thereby narrowing the gap to omnichannel applications.

The main contributions of this paper are: first, we create NonFoodKG with its ontologies based on Web information and link it to other Semantic Web sources. Second, we generalise environment information in such a way that it can be used by robot as well as Augmented Reality applications. Third, we showcase the omnichannel applicability of NonFoodKG in various shopping applications on robot, smartphone and smart glasses.

This work is structured as follows: In Section 2 we show how different kinds of knowledge were acquired for the use in NonFoodKG. Section 3 describes the components of the NonFoodKG knowledge graph with all ontologies. The omnichannel applicability of NonFoodKG is shown in Section 4 for example applications, Section 5 concludes.

2. 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 [23]. NonFoodKG consists of a location ontology created by accessing a robot map as well as modular product ontologies that contain structured or unstructured Web information as described in the following.

2.1. 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 as shown in [2]. We create a location ontology that is linked to precise environment information in a Digital Twin environment model created through robot stocktaking. A location ontology can be generated for any store.

The location ontology used in NonFoodKG is created using the knowledge processing framework KnowRob [17, 18]. Since KnowRob is an ontology-based system, the Digital Twin is stored in *Web Ontology Language (OWL)* format, allowing for more advanced reasoning on the comprised information. The *Web Ontology Language* defines a standard and vocabulary to process information content [24]. Knowledge in OWL format is described using triples of entities and their relations. The semantic Digital Twin therefore includes object properties, i.e. relations between individuals, in the following way: `<Product_Z> <has_pose> <Pose_A>`. It also includes data properties, i.e. relations of individuals to data like `<Pose_A> <has_xposition> <0.29>`.

The location ontology is based on the Digital Twin. The positions therein use a fixed reference frame with a given origin, usually set to a randomly chosen corner of the room that is to be scanned. 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 positions.

In the AR shopping applications we then use spatial perception[1] and match the world origin of the game to the reference frame origin of the Digital Twin. We also use world anchors[2] for locking shelf positions in the physical space, a technique that has proven effective in other work [25]. World anchors need to be created once and will be loaded in every subsequent run of the application. Since the location ontology stores positions of products in relation to shelf positions, product positions can also be inferred relative to these world anchors. Thus, digital content for specific products or shelves can be displayed.

In addition to the location ontology, we derive product dimension information from the perception system and for the dimension ontology. The robot uses its RGB and RGB-D sensors to identify the amount of consecutive products and calculate product width, height and depth.

2.2. Structured Web Information

NonFoodKG integrates and links to structured Web information. When structured data like ontologies are published on the Web and linked to each other, they are referred to as *Linked Data* [26], enabling machines to surf the Web. On the one hand, we integrate parts of structured Web information like the FoodKG ontologies [15] for a product-specific recipe recommendation. This is done by reusing the food product classification of the FoodOn ontology [13], which we intertwine with the product classification in the product taxonomy. NonFoodKG also integrates nutrition information from the Compositional Dietary Nutrition (cdno) ontology [27] as well as nutritional product information like Nutri-Score, product labels or packaging size for food products from Open Food Facts[3]. For food classes, we access product class nutrition information from the FoodData central [4] 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 NonFoodKG 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. [28]. 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 [29]. 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 Protegè[5]. Both the GO and the FoodOn workgroup are part of the OBO foundry [30], a community focussing on the development of interoperable ontologies for the biological sciences[6].

The product classification is linked to Wikidata[7] and the Product Types Ontology[8] wherever possible. We also use the Good Relations web vocabulary for E-Commerce[9]. Furthermore, NonFoodKG links to chemical information about allergens in the Chemical entities of biological interest (ChEBI)[10] ontology. The allergen ontology additionally links to other Semantic Web sources like the National Drug File (NDRF) of BioPortal[11] or Drugbank[12]. 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.

[1] Spatial perception in Unity: https://microsoft.github.io/MixedRealityToolkit-Unity/Documentation/SpatialAwareness/SpatialAwarenessGettingStarted.html
[2] World anchors in Unity: https://docs.unity3d.com/ScriptReference/XR.WSA.WorldAnchor.html
[3] Open Food Facts: https://world.openfoodfacts.org/
[4] Food Data central: https://fdc.nal.usda.gov/
[5] The Protegè website: https://protege.stanford.edu/
[6] obofoundry community: https://obofoundry.org/
[7] Wikidata: https://www.wikidata.org/wiki/Wikidata:Main_Page
[8] Product Types Ontology: http://www.productontology.org/
[9] Good Realtions vocabulary: http://www.heppnetz.de/ontologies/goodrelations/v1
[10] Chemical entities of biological interest: https://www.ebi.ac.uk/chebi/
[11] NDRF at BioPortal: https://bioportal.bioontology.org/ontologies/NDFRT/
[12] Drugbank: https://go.drugbank.com/

2.3. Unstructured Web Information

For creation of the NonFoodKG ontologies we acquire unstructured Web information to create structured knowledge sources that machines can query. In order for NonFoodKG to be usable in various environments, it needs a general product taxonomy that can be used for different retail domains like drugstores and grocery stores. Most techniques for taxonomy creation aim at properly modeling a domain by using a pattern-based approach for taxonomy induction [31] to learn simple but precise contextual patterns in a Web document. For a more generalised approach we use *Information Extraction* techniques on online store sitemaps of different retail domains to automatically create a more general product taxonomy. *Information Extraction* describes the process of converting text to structural information like entities and their relations [32].

Online stores are designed in a taxonomy-directed structure that can be used for Information Extraction. Previous attempts in taxonomy creation from online stores show its applicability but are spamming the used website to call children and sibling websites for Information Extraction [33, 34]. Since online stores usually also offer a sitemap in their robots file[13], 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[14] (grocery store) and dm[15] (drugstore). We then use Owlready [35] 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.

We additionally use Information Extraction to acquire product ingredient information from online stores for use in the ingredient ontology and information like brand, awarded labels, weight, filling capacity and country of origin from Codecheck[16], a consumer-oriented product information website. Furthermore, product ingredients are classified as allergens according to information from the U.S. Food & Drug administration (FDA) website[17].

3. Composition of NonFoodKG

Now that we have described the data acquired for the NonFoodKG product knowledge graph, we continue to show some metrics of NonFoodKG, explain the contained ontologies as well their accordance to the FAIR (Findable, Accessible, Interoperable, Reusable) principles [36] in the following.

3.1. Metrics

Table 1 depicts basic metrics of the NonFoodKG ontologies. These metrics are counts of axioms, classes, individuals and properties. Axioms are basic statements in ontologies that state hypotheses about what is true in a domain, whereas logical axioms secure *"the meaning of the logical connectives and quantifiers"* [37] in an ontology. Classes in ontologies are defined concepts that can contain subclasses, thereby forming a hierarchical classification - a taxonomy. Additionally, classes can contain instances: individuals. Please note that in this work we only use individuals for unique articles, which is different to approaches in other work. This means that we create a class for `thyme`, or `milk, 2.5% fat`, which can have various individuals assigned that have equal barcodes but differing best before dates, for example.

Object properties link two entities in an ontology, while data properties assign a datatype like a name or value to an individual. Annotation properties are used to annotate entities with additional information like language labels.

The product individuals and data, object and annotation properties are recurring in different ontologies, therefore their counts are not listed in the table for the overall NonFoodKG.

[13] A 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.

[14] Aldi-Nord sitemap: https://www.aldi-nord.de/.aldi-nord-sitemap.xml

[15] dm sitemap: https://www.dm.de/?view=asSitemap

[16] Codecheck website: https://www.codecheck.info/

[17] FDA website: https://www.fda.gov/

Ontology	axioms	logical axioms	classes	individuals	object properties	data properties	annotation properties
product-taxonomy	8,072	4,516	300	1,506	6	4	21
location ontology	7,721	6,715	7	978	3	8	8
dimension ontology	4,312	3,772	5	549	1	14	6
brand ontology	5,669	4,135	6	1,216	6	7	9
packaging ontology	5,435	4,786	14	574	7	4	6
label ontology	3,341	2,633	16	531	2	5	9
nutrition ontology	3,302	2,398	81	481	0	20	9
food-nutrition ontology	92,926	74,143	3,418	83	11	41	17
LinkFoodOn ontology	214	28	28	0	2	0	15
Non-food ingredient ontology	16,921	14,141	26	2,157	3	4	9
allergen ontology	892	300	33	197	3	2	11
NonFoodKG	148,805	117,567	3934				

Table 1: NonFoodKG ontology metrics

3.2. *Ontologies*

As depicted in Figure 2, the product builds the core of NonFoodKG. The modular ontologies in NonFoodKG have different foci that can be used to retrieve additional product information for different retail applications.

Product Taxonomy A customer entering a store to buy products from their shopping list will rather search for a product class (e.g."Where is the cereal?") than search for a specific product name. Other shopping applications try to solve this by remembering previously bought items [2] or offering a limited set of product categories that can be chosen from [38]. Using a product taxonomy to retrieve information about the product class offers some advantages: there is a broader range of products to be chosen from as well as the possibility to infer parent, sibling

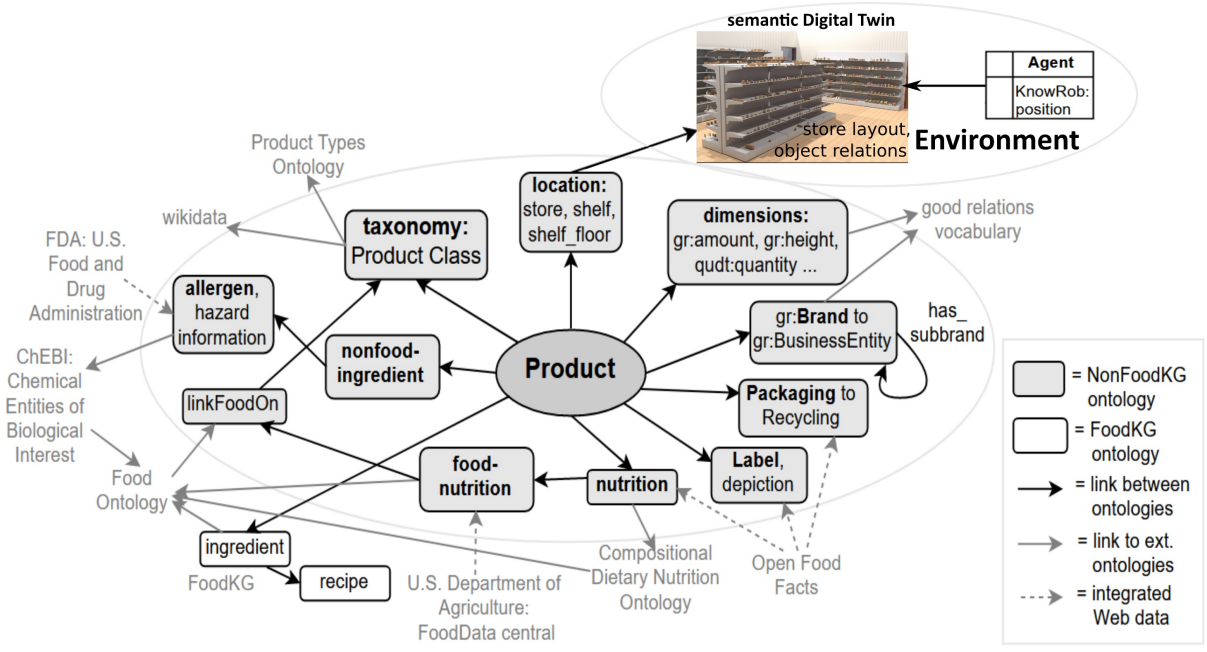

Fig. 2.: Overview of NonFoodKG ontologies and their links to external sources.

and children classes. Hence a shopping assistant can route a customer to the shelf holding cereal even if the store offers cereal in a broader context like `grain products`. What is more, if a customer is actually searching for a specific product from a certain brand which is not available in the given store, a product taxonomy can easily be used to retrieve parent or sibling class information to enable routing to the shelf with the next best solution or substitute products. Each product instance in the product taxonomy is assigned the Global Trade Identification Number (GTIN), a unique identifier encoded in the price tags of products. The entities of these product instances are reused in all other ontologies holding product-specific information.

Location Ontology One of the most commonly developed applications in shopping environments are routing applications as seen in [2, 39], for example. In [39], the difficulty of localisation of products in a store is shown. In contrast to this, in [40] a Digital Twin is successfully used to offer a standardised map with exact real-time article positions. We also use a Digital Twin store to create a location ontology that can additionally be used in Augmented Reality applications. This enables robots as well as smartphones, smart glasses or other devices to use current and exact location information for various applications as we demonstrate in this work.

Dimension Ontology For applications like refilling shelves, filling of new shelves or for fulfilment of click-and-collect orders, dimension information is needed. A sixpack of sponges might have the same dimensions as dishwasher detergent but weigh only 15 gram compared to 1500 gram. On a shelf holding detergent, only 8 articles might be placed next to each other, while on a shelf holding deodorants there might be 20 articles. Furthermore, plastic bags need to be grasped differently than carton boxes, which is important for robotic applications.

Brand Ontology Consumers are interested in product brand information and might search for store brands as substitutes for interesting products. They might as well be loyal to a certain brand and want to search for it. To allow for such applications, NonFoodKG includes the brand ontology, which classifies brands and allows for linking brands to a business entity.

Packaging Ontology The product packaging ontology is linked to a recycling ontology and therefore can be used to generate product recycling information for the product itself and all of its packaging material. Considering dishwashing detergent tablets in a 40 pieces carton as an example product, the tablet itself has hazardous ingredients and might be classified as residual waste. Since the tablets have their own plastic wrap (which we refer to as `inner_packaging`), this material is to be recycled as plastic waste. Lastly, the tablets come in an `outer_packaging`, a carton, that is to be recycled as paper waste. The packaging ontology can return all this information for a given product and can be used for recycling applications.

Label Ontology The product label ontology comprises information about labels that have been awarded to products. Labels certify certain quality standards of products, their ingredients, packaging or production processes and play an important role in shopping decision processes. All labels in the label ontology come with a description of the label criteria acquired from the Web and are categorized into label classes like `RecyclingLabel` for all labels certifying recycling standards. This enables a user to search for a label class in addition to a specific label. A consumer might want to search for all vegan products and not only products awarded the `V_label`, for instance. Supplementary, the label ontology is complemented with an image database to be used in AR applications.

Nutrition Ontology Nutrition ontologies play an important role in recommender systems as seen in [41, 42]. Since the nutrition ontology is comprised of different Web information, it includes both a nutrition classification and nutrition properties. The nutrition properties are being used for product-specific nutrition information, the nutrition classification is reused in the food-nutrition ontology.

Food-nutrition Ontology The food-nutrition ontology links the nutrition classification to product classes created from FoodData central information. In order for this information to be used in applications, it is also linked to the product taxonomy.

LinkFoodOn Ontology This ontology is extracted from the food product classification of the FoodOn ontology, which is then linked to from the food-nutrition ontology and intertwined with the product taxonomy. Thus, a product can be recognised in an application and its product-specific nutrition information can be queried from the nutrition ontology as well as its product class nutrition information from the food-nutrition ontology.

Non-food Ingredient Ontology For an amplification of semantic information, we use string matching to categorise ingredients. All ingredients enclosing the strings `alcohol`, `alcohols` or its german equivalent `alkohol` are categorized as `<alcohol> <rdf:type> <Alcohol>`, for example.

With the ingredients ontology a consumer can find all products that contain a certain ingredient or an ingredient class, like all facial care products containing acid, for example.

Allergen Ontology Many chemicals present in non-food products like cosmetics have been shown to cause contact dermatitis [43]. Therefore, the allergen ontology classifies ingredients as contact allergens.

While the ingredients ontology contains the information that `parfum` or `perfume` is a fragrance, the allergen ontology classifies `geraniol`, `linalool` or `limonene` as fragrances.

3.3. Accordance to FAIR principles

The FAIR guiding principles for scientific data management and stewardship [36] 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 NonFoodKG as explained in the following:

- Findability: NonFoodKG is openly accessible and uses persistent URLs (http://purl.org/NonFoodKG/)[18] for the contained ontologies that redirect the user to the current storage location of the ontologies.
- Accessibility: NonFoodKG offers different openly accessible endpoints for data access of robotic agents or humans (see Sec. 3.4).
- Interoperability: Section 2.2 describes the different vocabularies that are used or linked to in NonFoodKG. Hence, applications built on those ontologies can easily integrate NonFoodKG information.
- Reusability: Entities in the NonFoodKG ontologies use annotation properties like language labels, comments and links to term defining webpages such as wikidata or the product ontology as descriptors. All NonFoodKG ontologies come with a clearly stated licence, can be downloaded and reused.

3.4. Resource

NonFoodKG is publicly available at https://k4r-iai.github.io/NonFoodKG/ with all ontologies, a link to its triple store SPARQL endpoint[19] as well as an grlc[44] RESTful API for the example queries[20] demonstrated in the shopping assistant applications. We additionally provide a Sparklis [45] website[21] for users that are not familiar with SPARQL to access knowledge graph information.

4. Consumer Applications of NonFoodKG

Using NonFoodKG, we can answer complex questions about products, their ingredients, contained allergens, nutrition, labels, brands, dimensions, packaging or location. The triple store used for storage of NonFoodKG offers a *SPARQL* endpoint for information access. *SPARQL* is the standard query language for triple data. We continue to discuss the communication of Unity game engine and KnowRob processing system with the SPARQL endpoint for the AR shopping assistant and the robot shopping assistant, respectively.

4.1. Communication

Both Augmented Reality and robot shopping assistants can use the same SPARQL queries for shopping applications. Since robot assistants encode their environments differently (as discussed in 2.1), they need to additionally relate the query results to the Digital Twin store.

[18] Link to the purl administration page: https://purl.prod.archive.org/

[19] The NonFoodKG API can be accessed at this SPARQL endpoint: https://api.krr.triply.cc/datasets/mkumpel/FoodToNonFoodKG/services/FoodToNonFood/sparql

[20] Example queries can be accessed online: http://grlc.io/api/K4R-IAI/NonFoodKG/SPARQLfiles/

[21] The NonFoodKG Sparklis website can be accessed here: http://nonfoodkg.informatik.uni-bremen.de/

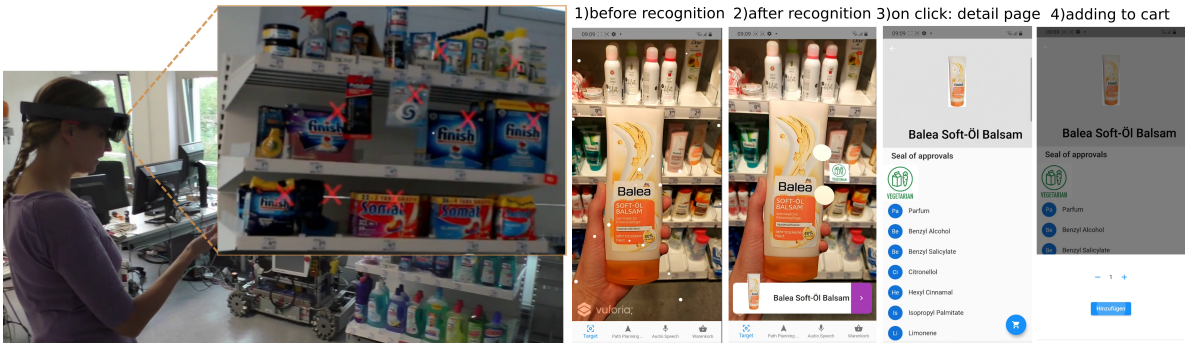

Fig. 3.: Left: Preference visualisation using HoloLens. Right: Visualising product information on smartphone.

Augmented Reality Shopping Assistant The Unity game development platform offers a library for making Web requests called UnityWebRequest[22] as well as a JSON Serialization to convert Unity objects to and from JSON format[23], which we use for processing of the SPARQL results.

Robot Shopping Assistant One of the benefits of using NonFoodKG as a single point of access is its platform-independency. Thus a robotic agent running KnowRob can easily connect product to environment information in a Digital Twin as well. KnowRob uses the logic programming language Prolog[24] for action execution. Prolog offers a SPARQL client library that can be used to process SPARQL query results.

4.2. *Example Augmented Reality Applications*

The following applications demonstrate the applicability of the NonFoodKG knowledge graph as one source of data for a range of consumer applications.

Application 1: Preference Visualisation in AR using HoloLens This openly accessible application[25] highlights all objects that do not comply with a set preference on a HoloLens. Due to the fact that digital overlays hinder consumers to look at the products behind the overlay, we decided to highlight opposing products. Thus, for a consumer searching for organic products, all non-organic products will be highlighted by overlaying a digital, red `X`. By highlighting all opposing products, the customer can further examine all products that do comply with their preference.

To localize the HoloLens in the retail laboratory we use spatial perception[26] and match the world origin of the Unity game to the reference frame origin of the robot environment map as discussed in 2.1. Product positions are inferred relative to the existing world anchors, which are attached to shelf instances. Each product is created as a `X` game object named according to the NonFoodKG product names. Thereby, objects can be set visible or invisible 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 3.

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 with the Vuforia Engine [27]. Therefore, we rely on images or barcodes of products available at the retail laboratory.

---

[22] Unity web requests library: https://docs.unity3d.com/Manual/UnityWebRequest.html
[23] JSON serialization in Unity: https://docs.unity3d.com/Manual/JSONSerialization.html
[24] SWI-Prolog: ttps://www.swi-prolog.org/
[25] HoloLens application: https://github.com/michaelakuempel/HoloPreferenceDemo
[26] https://microsoft.github.io/MixedRealityToolkit-Unity/Documentation/SpatialAwareness/SpatialAwarenessGettingStarted.html
[27] Image targets with Vuforia: https://library.vuforia.com/features/images/image-targets.html

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 are 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 3. It shows four phases of the application: 1) the customer is browsing the aisles, no object has been recognised 2) an object is recognised, all product label images appear next to the product and the product name is displayed in an information slider at the bottom 3) a product detail page is being displayed 4) the product can be added to the shopping cart in the desired quantity.

4.3. Device-independent 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 routing applications.

Application 3: AR routing to product destination using HoloLens or smartphone This shopping application can be applied to various Augmented Reality devices, which we show by deploying it to the HoloLens and a smartphone. In the routing application, a customer is led to a searched product through the display of digital waypoints as displayed in Figures 1 and 4. The application uses spatial perception and world anchors as described in Sec. 4.2 to lock digital content to real world positions. It also uses a navigation mesh[28] to define obstacles and paths to shelves in the Digital Twin map. Digital waypoints to product destinations are aligned relative to the world anchors and designed in such a way that they constantly adapt to changes of device direction. This means that a customer can start walking into the opposite direction and the waypoints will be updated to lead them to the target product. For the solution in this work, all product categories of one shelf belong to the same destination point. Using Unity, the project can easily be exported for use on different platforms.

Figure 1, middle, shows a speech interface that can be used for routing applications, but the applications additionally provide a website menu (see Fig. 4, left) that automatically creates a list of available product categories from the accessed knowledge graph[29]. Figure 4, middle, then shows the application for a customer searching for `deodorant`. In the HoloLens application (Fig. 1), the shelve of the searched product is being highlighted by an orange destination point while green waypoints are augmented onto the path to lead the customer to the searched product category. In the smartphone application, instead of waypoints we display the route as a line and include a mini-map of the store for better orientation.

[28] Navigation mesh in Unity: https://docs.unity3d.com/ScriptReference/AI.NavMesh.html

[29] The idea of creating a website from a knowledge graph is inspired by: http://matthewedward.blogspot.com/2014/04/create-html-web-pages-linked-data-php-sparql-arc2.html

Fig. 4.: Routing Applications: Left: Website that can be displayed on Ar device or robot tablet. Middle: AR routing to product destination using Smartphone. Right: Robot leading a customer to a searched product.

Application 4: Robot routing to product destination Robots can use NonFoodKG by accessing it for product information to then link the retrieved information to exact environment information in the Digital Twin. Therefore, the robot running the KnowRob knowledge processing system will query NonFoodKG in Prolog as shown in Figure 5. The prefix declarations for this query can be found in Table 2. The Prolog query calls a similar SPARQL query from a defined SPARQL endpoint as the AR routing applications described previously. The SPARQL query in lines 2 to 7 returns the product GTIN of all searched products. In this example, the query searches for all **`detergents`** that are on **`sale`**. In case of more than one results, the result set is ordered by ascending price. Afterwards (lines 9 to 11), the returned product GTIN is matched to the Digital Twin representation to retrieve the object position as `T` transformation and `R` rotation in relation to the environment map origin. The application is depicted in Fig. 4, right.

```
1 sparql_query(                                  # call SPARQL query
2   'select ?GTIN where                            # return GTIN variable
3      ?product rdf:type tax:detergent.            # retrieve all products of type detergent
4      ?product gr:hasEAN_UCC-13 ?GTIN.            # retrieve GTIN variable
5      ?product loc:has_flag loc:"SALE".           # for all products that are on sale
6      ?instance loc:has_price ?price.             # retrieve price for sorting ascending
7    } ORDER BY ?price'
8    , Row, [ endpoint('SPARQL endpoint')]),   # from SPARQL endpoint
9  get_product_type(?GTIN, P),                    # retrieve internal product ID to GTIN
10 has_type(I, P),                              # retrieve object item to product ID
11 is_at(I, ['map', T, R]).                     # retrieve position for product
```

Fig. 5.: Example Prolog query to retrieve the exact position of detergent that is on sale.

4.4. Generalised Information For Personalised Applications

NonFoodKG offers a basic user profile ontology that can be integrated and instantiated in ontologies like the food-nutrition, label or non-food ingredient ontology for personalised applications. Figure 6 shows an example product recommender application based on personal dietary preferences[30]. 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

[30]The Web Applications are available at: http://nonfoodkg.informatik.uni-bremen.de/

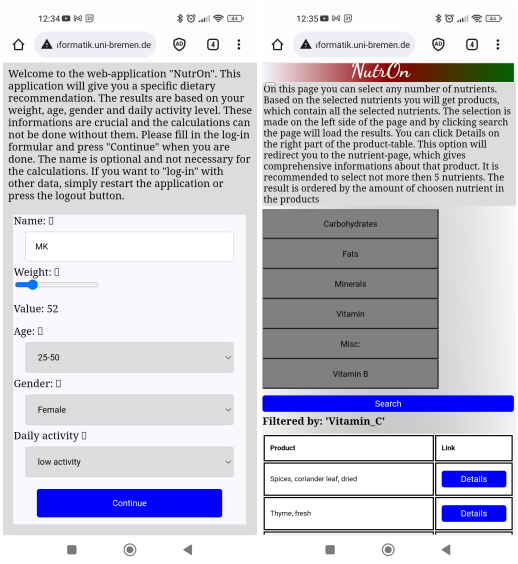

Fig. 6.: Product recommender for nutritional needs.

Prefix	IRI
rdf	http://w3.org/1999/02/22-rdf-syntax-ns#
tax	http://purl.org/NonFoodKG/product-taxonomy
gr	http://purl.org/goodrelations/v1
loc	http://purl.org/NonFoodKG/location

Table 2: Prefixes used in this work and their corresponding namespaces

on the given input values. In this example, a female user is searching for products that are rich in vitamin c. The results page shows products in accordance to the preferences and query.

The user profile can similarly be used in other domains to set preferences like allergies or ecological and economical interests that can be applied to the shopping applications shown in Figure 3.

In addition to this personalisation, NonFoodKG offers generalised knowledge through its modularity. Consider a person planning dinner. They might also be planning on going running in the afternoon and therefore want to cook a meal that contains plenty of magnesium. This person might use the previously introduced app for a product recommendation for their dietary preferences to then use NonFoodKG in combination with the connected FoodKG[31] to query for recipes that contain these products. Through its connection to current Digital Twin information, the user can also access NonFoodKG to check if the products are actually available in the store. While in the store, NonFoodKG can be accessed to help the user find the products as shown in the routing applications. Lastly, when preparing dinner, the user can access NonFoodKG to open the recipe website for cooking instructions.

5. Discussion

This work describes the NonFoodKG knowledge graph with all its ontologies and points out the benefits of using a knowledge graph as a single access point with set standards for integration of different knowledge sources. 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 NonFoodKG by HoloLens, smartphone and robot. 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. NonFoodKG offers:

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

With this introduction of NonFoodKG we hope to start discussions about a standard model for non-food products that links to domains like taxonomy, chemical composition, diseases, product brands and labels or ingredient origin, amongst others. As discussed previously, NonFoodKG reuses many terms from other ontologies to provide a standardised vocabulary. While we connect NonFoodKG to a Digital Twin store model with product information, online stores can similarly access the contained information. However, the current version of NonFoodKG contains only 949 distinct product individuals for an example store and does not yet provide an automated information gathering feature for new individuals in a linked Digital Twin. Websites change almost as often as offered products in a store, which poses another challenge in automated content creation that we want to focus on in future work.

Additionally, besides consumers trusting some agents more than others, they also trust some information sources more than others. This is especially true for Web information. What is more, a robot needs to trust the information source to successfully and safely operate in its environment. Therefore, enabling users to only access trusted sources in the different applications is planned for a future release.

5.1. Sustainability Plan and Future Work

The NonFoodKG knowledge graph has been developed as part of the Knowledge4Retail project[32] and we expect to support and expand NonFoodKG actively for the next 3-8 years. We plan on expanding the ontologies to integrate

[31] FoodKG: https://foodkg.github.io/
[32] Knowledge4Retail project website: https://knowledge4retail.org/

even more knowledge sources and to develop more robotic applications that can embed knowledge graph information. For this, we will integrate a more seamless logic SPARQL programming language in Prolog: sparqlprog[33]. Additionally, we plan to integrate recipe information, expand the allergen and disease ontology and include a symptom ontology. All applications are under development and will be updated over time.

References

[1] G. Waltner, M. Schwarz, S. Ladstätter, A. Weber, P. Luley, H. Bischof, M. Lindschinger, I. Schmid and L. Paletta, MANGO-mobile augmented reality with functional eating guidance and food awareness, in: *International Conference on Image Analysis and Processing*, Springer, 2015, pp. 425–432.

[2] Z. Davis, M. Hu, S. Prasad, M. Schuricht, P. Melliar-Smith and L.E. Moser, A personal handheld multi-modal shopping assistant, in: *International conference on Networking and Services (ICNS'06)*, IEEE, 2006, pp. 117–117.

[3] H.-M. Gross, H.-J. Boehme, C. Schröter, S. Müller, A. König, C. Martin, M. Merten and A. Bley, Shopbot: Progress in developing an interactive mobile shopping assistant for everyday use, in: *2008 IEEE International Conference on Systems, Man and Cybernetics*, IEEE, 2008, pp. 3471–3478.

[4] Y. Chen, F. Wu, W. Shuai and X. Chen, Robots serve humans in public places—KeJia robot as a shopping assistant, *International Journal of Advanced Robotic Systems* **14**(3) (2017), 1729881417703569.

[5] R. Caccavale, P. Arpenti, G. Paduano, A. Fontanellli, V. Lippiello, L. Villani and B. Siciliano, A flexible robotic depalletizing system for supermarket logistics, *IEEE Robotics and Automation Letters* **5**(3) (2020), 4471–4476.

[6] M. Costanzo, G. De Maria, G. Lettera and C. Natale, Can robots refill a supermarket shelf?: Motion planning and grasp control, *IEEE Robotics & Automation Magazine* **28**(2) (2021), 61–73.

[7] W. Lockie, Delivering an effective click-and-collect strategy: A retailer case study, *Journal of Digital & Social Media Marketing* **2**(2) (2014), 139–152.

[8] A.-S. Riegger, K. Merfeld, J.F. Klein and S. Henkel, Technology-enabled personalization: Impact of smart technology choice on consumer shopping behavior, *Technological Forecasting and Social Change* **181** (2022), 121752.

[9] D. Taylor, S. Brockhaus, A.M. Knemeyer and P. Murphy, Omnichannel fulfillment strategies: defining the concept and building an agenda for future inquiry, *The International Journal of Logistics Management* (2019).

[10] A. Albadvi and M. Shahbazi, A hybrid recommendation technique based on product category attributes, *Expert Systems with Applications* **36**(9) (2009), 11480–11488.

[11] T. Berners-Lee, J. Hendler, O. Lassila et al., The semantic web, *Scientific american* **284**(5) (2001), 28–37.

[12] E. Sirin, J. Hendler and B. Parsia, Semi-automatic composition of web services using semantic descriptions, in: *1st Workshop on Web Services: Modeling, Architecture and Infrastructure*, 2003, pp. 17–24.

[13] D.M. Dooley, E.J. Griffiths, G.S. Gosal, P.L. Buttigieg, R. Hoehndorf, M.C. Lange, L.M. Schriml, F.S. Brinkman and W.W. Hsiao, FoodOn: a harmonized food ontology to increase global food traceability, quality control and data integration, *npj Science of Food* **2**(1) (2018), 1–10.

[14] J. Cantais, D. Dominguez, V. Gigante, L. Laera and 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.

[15] S. Haussmann, O. Seneviratne, Y. Chen, Y. Ne'eman, J. Codella, C.-H. Chen, D.L. McGuinness and M.J. Zaki, FoodKG: a semantics-driven knowledge graph for food recommendation, in: *International Semantic Web Conference*, Springer, 2019, pp. 146–162.

[16] M.N.K. Boulos, A. Yassine, S. Shirmohammadi, C.S. Namahoot and M. Brückner, Towards an "Internet of Food": food ontologies for the internet of things, *Future Internet* **7**(4) (2015), 372–392.

[17] M. Tenorth and M. Beetz, KnowRob—knowledge processing for autonomous personal robots, in: *2009 IEEE/RSJ international conference on intelligent robots and systems*, IEEE, 2009, pp. 4261–4266.

[18] M. Beetz, D. Beßler, A. Haidu, M. Pomarlan, A.K. Bozcuoğlu and G. Bartels, KnowRob 2.0—a 2nd generation knowledge processing framework for cognition-enabled robotic agents, in: *2018 IEEE International Conference on Robotics and Automation (ICRA)*, IEEE, 2018, pp. 512–519.

[19] D. Beßler, R. Porzel, M. Pomarlan, A. Vyas, S. Höffner, M. Beetz, R. Malaka and J. Bateman, Foundations of the socio-physical model of activities (SOMA) for autonomous robotic agents, *arXiv preprint arXiv:2011.11972* (2020).

[20] M. Kümpel, C.A. Müller and M. Beetz, *Semantic Digital Twins for Retail Logistics*, in: *Dynamics in Logistics: Twenty-Five Years of Interdisciplinary Logistics Research in Bremen, Germany*, M. Freitag, H. Kotzab and N. Megow, eds, Springer Nature Switzerland AG, Gewerbestrasse 11, 6330 Cham, Switzerland, 2021, p. ?–?.

[21] M. Grieves, *Virtually perfect: driving innovative and lean products through product lifecycle management*, Space Coast Press, 2011.

[22] P. Augustine, The industry use cases for the digital twin idea, in: *Advances in Computers*, Vol. 117, Elsevier, 2020, pp. 79–105.

[23] N. Heist, S. Hertling, D. Ringler and H. Paulheim, Knowledge Graphs on the Web–An Overview, *Knowledge Graphs for eXplainable Artificial Intelligence: Foundations, Applications and Challenges* (2020), 3–22.

[24] D.L. McGuinness, F. Van Harmelen et al., OWL web ontology language overview, *W3C recommendation* **10**(10) (2004), 2004.

[33] sparqlprog on SWI-Prolog: https://www.swi-prolog.org/pack/list?p=sparqlprog

[25] A. Jakl, L. Schöffer, M. Husinsky and M. Wagner, Augmented Reality for Industry 4.0: Architecture and User Experience., in: *FMT*, 2018, pp. 38–42.
[26] C. Bizer, T. Heath and T. Berners-Lee, Linked data: The story so far, in: *Semantic services, interoperability and web applications: emerging concepts*, IGI global, 2011, pp. 205–227.
[27] L. Andrés-Hernández, A. Baten, R. Azman Halimi, R. Walls and G.J. King, Knowledge representation and data sharing to unlock crop variation for nutritional food security, *Crop Science* **60**(2) (2020), 516–529.
[28] C. Bizer, R. Cyganiak, T. Heath et al., How to publish linked data on the web (2007).
[29] G.O. Consortium, The gene ontology resource: 20 years and still GOing strong, *Nucleic acids research* **47**(D1) (2019), D330–D338.
[30] 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**(11) (2007), 1251–1255.
[31] H. Davulcu, S. Koduri and 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.
[32] O. Etzioni, M. Banko, S. Soderland and D.S. Weld, Open Information Extraction from the Web, *Commun. ACM* **51**(12) (2008), 68–74. doi:10.1145/1409360.1409378.
[33] H. Davulcu, S. Vadrevu, S. Nagarajan and I. Ramakrishnan, OntoMiner: Bootstrapping and populating ontologies from domain-specific Web sites, *Intelligent Systems, IEEE* **18** (2003), 24–33. doi:10.1109/MIS.2003.1234766.
[34] M. Tenorth, U. Klank, D. Pangercic and M. Beetz, Web-Enabled Robots, *IEEE Robotics Automation Magazine* **18**(2) (2011), 58–68. doi:10.1109/MRA.2011.940993.
[35] J.-B. Lamy, Owlready: Ontology-oriented programming in Python with automatic classification and high level constructs for biomedical ontologies, *Artificial intelligence in medicine* **80** (2017), 11–28.
[36] 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**(1) (2016), 1–9.
[37] J.S. Tanton, *Encyclopedia of mathematics*, Infobase Publishing, 2005.
[38] C. Thompson, H. Khan, D. Dworakowski, K. Harrigan and G. Nejat, An Autonomous Shopping Assistance Robot for Grocery Stores, in: *Workshop on Robotic Co-workers 4.0: Human Safety and Comfort in Human-Robot Interactive Social Environments 2018 IEEE/RSJ*.
[39] N. Doering, S. Poeschl, H.-M. Gross, A. Bley, C. Martin and H.-J. Boehme, User-centered design and evaluation of a mobile shopping robot, *International Journal of Social Robotics* **7**(2) (2015), 203–225.
[40] B. Brenner and V. Hummel, Digital twin as enabler for an innovative digital shopfloor management system in the ESB Logistics Learning Factory at Reutlingen-University, *Procedia Manufacturing* **9** (2017), 198–205.
[41] V. Espín, M.V. Hurtado and M. Noguera, Nutrition for Elder Care: a nutritional semantic recommender system for the elderly, *Expert Systems* **33**(2) (2016), 201–210.
[42] P. Tumnark, F.A.d. Conceição, J.P. Vilas-Boas, L. Oliveira, P. Cardoso, J. Cabral and N. Santibutr, Ontology-based personalized dietary recommendation for weightlifting (2013).
[43] I. Kimber, D.A. Basketter, G.F. Gerberick and R.J. Dearman, Allergic contact dermatitis, *International immunopharmacology* **2**(2–3) (2002), 201–211.
[44] A. Meroño-Peñuela and 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. ISBN 978-3-319-47602-5. doi:10.1007/978-3-319-47602-5_48.
[45] S. Ferré, Sparklis: An expressive query builder for SPARQL endpoints with guidance in natural language, *Semantic Web* **8**(3) (2017), 405–418.