



# TaxoMulti: Rule-Based Expert System to Customize Product Taxonomies for Multi-Channel E-commerce

Heiko Angermann<sup>1</sup>

Received: 1 November 2021 / Accepted: 12 February 2022 / Published online: 1 March 2022  
© The Author(s) 2022

## Abstract

Taxonomies are used as product categories to facilitate users navigating through an e-commerce portal with the help of hierarchically structured concepts. However, the identical taxonomy is shown to each customer regardless of the channel used. This is challenging for the customers in terms of user experience, as the screen size is rigid, and has not a flexible format like a printed catalog. Simply reducing the taxonomy as suggested in existing works is not sufficient, as it leads to semantic misrepresentation of the product domain. To overcome the inflexibility of product taxonomies, the rule-based expert system *TaxoMulti* is presented in this paper. The main objectives of our descriptive research are the formulation of the taxonomy over- and undersize problem in multi-channel context, before different types of flexible mediator concepts are discussed that allow overcoming these challenges. Using our novel method, marketing experts can now provide different taxonomies including the same semantics to be shown on different channels. The method is implemented using logic programming, allowing the integration of an inference engine utilizing background knowledge without changing the underlying logic of the used information (management) system. The comprehensive experiments on three public and private databases highlight the improvement when adding different types of mediator concepts for the adaption process. Compared to existing best performing works in related fields, *TaxoMulti* has achieved an improvement of + 26.31 % for the reduction of the taxonomy, + 60 % for the enlargement of the taxonomy, and + 21.21 % in terms of flexibility.

**Keywords** Rule-based system · Expert system · Rule-based publishing · Taxonomy · Multi-channel · Web engineering

## Introduction

During the last decade, the use of online portals has increased significantly [14, 37]. The main reason is the emergence of mobile devices, leading to various channels, known as multi-channel paradigm, or as omnichannel if the multi-channels are connected with each other [9, 26]. Providing a multi-channel e-commerce portal improves user experience, but it is complex due to its diversity [16, 27, 30, 45, 51]. This is because the online channels (e.g. desktop, mobile) are based on various device classes. Each differs according to its in-/output capabilities, especially its screen size [38, 52, 59].

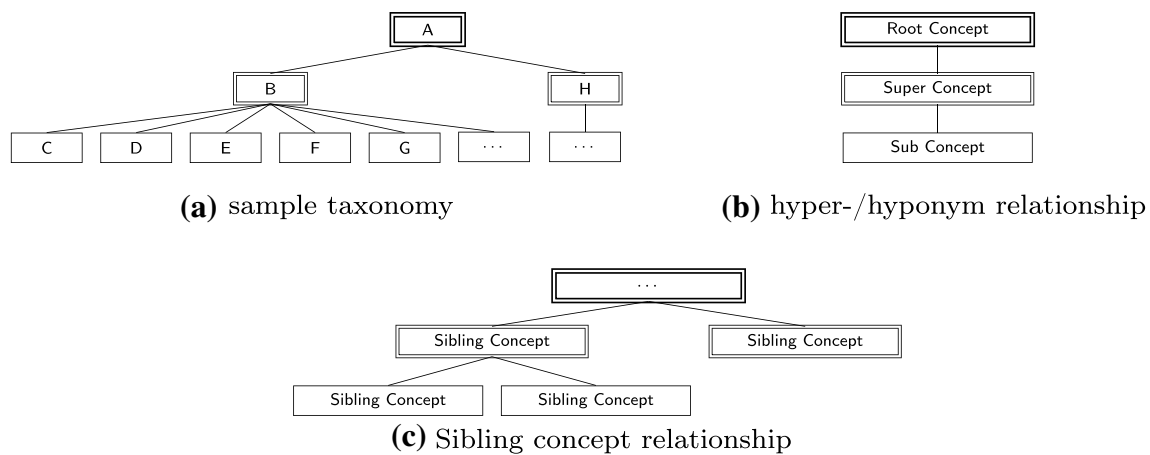
Regardless of the number of different channels provided, taxonomies have proven to be an indispensable method for

operating information management systems like e-commerce online portals [43, 44, 48]. The taxonomy is used in the backend of the application to formally structure product categories in a hierarchical manner [20]. In the frontend, it is displayed to the users to support exploring product categories. However, regardless of the domain, when implementing a multi-channel portal, the focus is merely on channel-specific technical perspectives. For example, different screen sizes of device classes are considered using style sheets. The channel-specific capabilities are not considered from a semantic perspective. The taxonomy shown to the user is identical, regardless of the channel utilized. This is because the creation of the taxonomy is formal and often historically grown, as most retailers use the taxonomy additionally in other entrepreneurial perspectives (e.g. product master data, media asset management) [10, 61].

Existing research paradigms aiming to customize product taxonomies are catalog segmentation, dynamic taxonomies, and personalized directories [7]. The systems for catalog segmentation are assigning the users to a pre-defined

✉ Heiko Angermann  
heiko.angermann@h-da.de

<sup>1</sup> Faculty of Design, Hochschule Darmstadt University of Applied Sciences, Olbrichweg 10, 64287 Darmstadt, Germany



**Fig. 1** The hierarchical structure of a sample taxonomy

expert-generated sub taxonomy, which forms a sub catalog. Catalog segmentation played a crucial role in commercial printing. For example, a furniture retailer publishes a main catalog once a year, and provides additional sub catalogs for different seasons [12]. The latest works on catalog segmentation are using preference techniques as used in recommender systems to automate allocating users to the most related sub taxonomy (sub catalog), e.g. in [36, 37, 39, 52, 63]. However, the sub taxonomies are semantically reduced, which means that concepts not included in the sub catalog are not shown. However, the needs regarding furniture are different depending on the region, the age of the customer, and of course the personal preferences. This is a main drawback, as additionally, the preferences change over time. Techniques of personalized directories are overcoming most of the before-mentioned problems, as those are aiming to modify the taxonomy according to users' personal requirements, e.g. in [6, 21, 22, 28, 34, 47, 62]. The most recent technique is providing various modification rules on subsets of concepts, which allows to not semantically reduce the taxonomy. However, the flexibility of the rules provided still lacks in sufficient support for multi-channel. Each user has indeed a personalized taxonomy, but it is identical across channels. In multi-channel context, some users are single-channel users, but most are multi-channel users [13, 45, 60]. The third related research paradigm named dynamic taxonomies, prunes the taxonomy according to the informal keyword given manually by the user, e.g. in [15, 24, 29, 31, 50, 56, 58]. Using dynamic taxonomies, the user can most directly interact with the online portal, but the domain is semantically reduced according to the keyword.

To provide a system to effectively create flexible taxonomies in multi-channel context, *TaxoMulti* is presented. The core motivation of the proposed system is to overcome the taxonomy oversize and taxonomy undersize challenge without suffering from semantic misinterpretation after the

adaption process. This is achieved through studying two novel types of mediator concepts that are flexible and created as background knowledge resources. The contributions are as follows:

- It is the first paper that discusses the problem of using taxonomies in multi-channel context from a semantic perspective. This is achieved by formulating the significance of the problems defined as taxonomy undersize and oversize.
- It provides a novel method for creating flexible taxonomies by using a dynamic adaption process. This is achieved by defining different types of latent mediator concepts and various modification rules being stored as background knowledge.
- It implements the method using an architecture that does not change the underlying initial logic of the information management system. This is achieved by implementing the method using a system independent inference engine.

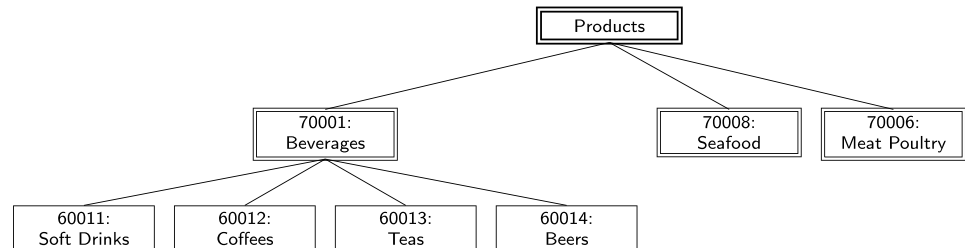
The rest of the paper is organized as follows. Section “[Background](#)” formulates and discusses the taxonomy problem in a multi-channel context. The proposed rule-based expert system *TaxoMulti* is presented in section “[Proposed system TaxoMulti](#)”. A case study for three different channels is presented in section “[Case study and demonstration](#)”. In section “[Experimental evaluation](#)”, the proposed system is evaluated using three databases representing retailing markets. The work concludes in section “[Conclusions](#)”.

## Background

A *Taxonomy* ( $\Theta$ ), as given in Fig. 1a, is a hierarchy of objects with similar properties [7], defined as technical terms representing domains [46, 49] with (see Eq. 1):

**Table 1** Comparison between taxonomy and ontology regarding semantic relationships

Symbol	Meaning	Relationship	Example	Ontology	Taxonomy
$\supset$	Super-ordinate	Hypernym	“vehicle” $\supset$ “car”	Yes	Yes
$\subset$	Sub-ordinate	Hyponym	“car” $\subset$ “vehicle”	Yes	Yes
$\in$	Part-of	Meronym	“wheels” $\in$ “car”	Yes	No
$\neq$	Opposite	Antonym	“move” $\neq$ “stand”	Yes	No
$=$	Identical	Synonym	“car” = “auto”	Yes	No

**Fig. 2** A taxonomy (excerpt) as provided by the public *Northwind* database

$$\Theta = (\{\Phi\}, \{A\}), \quad (1)$$

where  $\Phi$  is a partially ordered set of concepts, and  $A$  is a set of edges. Each edge represents a hierarchical relationship between single concepts in an acyclic graph (see Fig. 1b), referred as *is-a* relationships [19]. Each concept is defined by a name, i.e. the label. Using the label and the taxonomic path resulting through the *is-a* relationships, each concept represents a more or less generalized and in turn fine-granular semantic excerpt of the entire domain [7]. Formally, the above-mentioned *is-a* relationship are known as hypernym and hyponym [19]. In contrast to an *Ontology*, a taxonomy is only limited to this semantic relationships (summarized in Table 1) [25, 57]. Whereas the ontologies are used for knowledge management (e.g. Wordnet [23]), the taxonomies are used for information management, as the *is-a* relationships are mostly sufficient.

### Concept Types

Depending on the vertical and horizontal position of the concept, different concept types are distinguished (see Fig. 1b): sub concept, super concept, root concept, and sibling concept. A concept named *Sub Concept*, formally *subof*, is a less generalized concept that sub-ordinates another one, as given in Eq. 2, if:

$$D = \text{subof}(B) : \Leftrightarrow (D \subset B) \wedge ((D \wedge B) \in \Theta), \quad (2)$$

where  $D$  and  $B$  are two concepts of taxonomy  $\Theta$ .

A concept named *Super Concept*, formally *superof*, is a more generalized concept that super-ordinates another concept, as given in Eq. 3, if:

$$B = \text{superof}(D) : \Leftrightarrow D = \text{subof}(B), \quad (3)$$

where  $D$  and  $B$  are two concepts of taxonomy  $\Theta$ .

A *Root Concept*  $A$ , formally *rootof*, is the most general concept that is not super-ordinated by another concept, as given in Eq. 4, in which:

$$A = \text{rootof}(\Theta) : \Leftrightarrow \nexists \text{superof}(A). \quad (4)$$

where  $A$  is a concept of taxonomy  $\Theta$ .

A concept named *Sibling Concept*, formally *sibof*, is the relationship between two concepts sharing the identical super-ordinated concept and have the same horizontal position, as given in Eq. 5, if (see Fig. 1c):

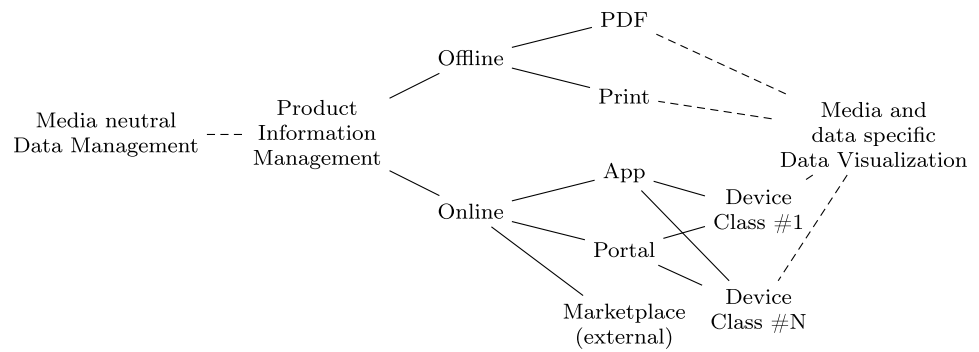
$$E = \text{sibof}(D) : \Leftrightarrow (E \wedge D) = \text{subof}(B). \quad (5)$$

where  $E$  and  $D$  are sub concepts of  $B$ , which are all concepts of taxonomy  $\Theta$ .

Depending on how many vertical edges inside the graph exist that are not sibling relationships, the taxonomy consists of various *Levels*. Our example taxonomy presented below consists of three levels (see Fig. 2):

1. The concept “Products” is the root concept.
2. The concepts “Seafood”, “Meat Poultry”, and “Beverages” sub-ordinate the root concept, and are sibling concepts. “Products” is their common super concept.
3. The concepts “Soft Drinks”, “Coffees”, “Teas”, and “Beers” sub-ordinate the concept “Beverages”, and are sibling concepts. Each is-a sub concept of the “Beverages”, which is their super concept. The concepts “Soft Drinks”, “Coffees”, “Teas”, and “Beers” do not super-ordinate another concept(s).

**Fig. 3** Using a product information management system for crossmedia multi-channel



## Problem Formulation

Taxonomies are a formal way of representing information, and are utilized for hierarchically structuring data in (enterprise) information management systems [18], e.g. enterprise resource planning, online portals, document management systems [2, 3, 5, 18, 55]. The taxonomies are generated through an expert, or by referring to a standard taxonomy, e.g. the *North American Product Classification System*<sup>1</sup> (NAPCS) [17, 53]. In e-commerce, the online portal application itself mainly acts as shopping cart for representing and ordering products. The actual product information is very often maintained in another system - often in a product information management (PIM) system. The necessary data enters the e-commerce database (backend) through an interface with the PIM or an additionally system referred as middleware operating between the PIM and the shopping cart [35]. Usually, the PIM has interfaces to further applications, as the PIM system stores the information central and media neutral. This is important when the information is published in crossmedia context, meaning that the media must be visualized on multiple specific online as well as offline channels [42], but the data must be consistent across all different channels (see Fig. 3).

As every single concept represents a more or less generalized semantic excerpt of the entire domain, modifying the taxonomy happens rarely [7]. Removing, adding, but also modifying (label or level) a single concept would effect that the information gain changes and the entire semantic would be misleading [54]. This intentional formality of taxonomies results in inflexibility. And, it also has significant drawbacks when the taxonomy is published in the frontend of multi-channel e-commerce portals:

1. As for channels using smaller device classes, the full range of concepts can't be displayed in a user-friendly manner, it suffers from a problem which in the following is called *Taxonomy Oversize*. Instead of requiring

massive input from the user to navigate to the concept desired (e.g. scrolling, tapping/clicks, filtering), the taxonomy should be reduced by a number of concepts  $Y$ , but without losing any information about the domain, indicated with (see Eq. 6):

$$Y = |\Phi_{\alpha} - \Phi_{\beta}|, \quad (6)$$

where  $\Phi_{\alpha}$  is the set of concepts detailing the initial taxonomy, and  $\Phi_{\beta}$  includes the concepts detailing the reduced taxonomy. As taxonomies usually consist of multiple levels, the desired reduction must be considered according to the level of the concepts. For sub concepts, the desired taxonomy reduction  $Y_{sub}$  can be formulated as (see Eq. 7):

$$Y_{sub} = |\Phi_{\alpha_{sub}} - \Phi_{\beta_{sub}}|, \quad (7)$$

where  $\Phi_{\alpha_{sub}}$  is the set of sub concepts detailing the initial taxonomy, and  $\Phi_{\beta_{sub}}$  includes the sub concepts detailing the reduced one. Accordingly, for super concepts, the desired reduction  $Y_{super}$  can be evaluated (see Eq. 8):

$$Y_{super} = |\Phi_{\alpha_{super}} - \Phi_{\beta_{super}}|, \quad (8)$$

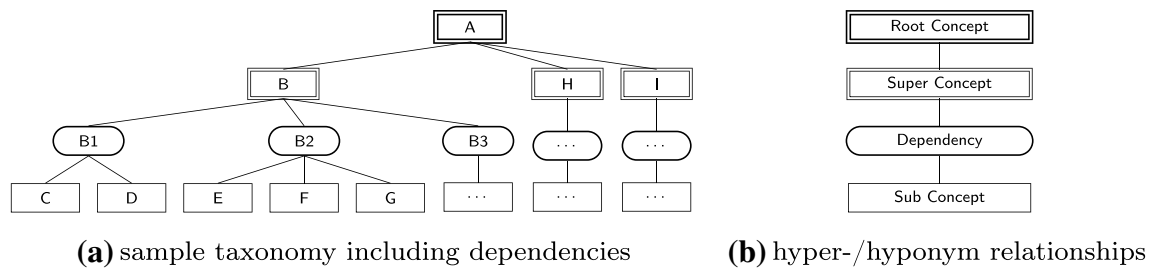
where  $\Phi_{\alpha_{super}}$  is the set of super concepts detailing the initial taxonomy, and  $\Phi_{\beta_{super}}$  only includes the super concepts detailing the reduced one. Note that both equations can of course be modified according to further levels.

2. As for channels using larger device classes, the range of concepts should be more detailed to help users more quickly finding the desired products, it suffers from a problem which in the following is called *Taxonomy Oversize*. Instead of requiring a good understanding of the domain by the user, the taxonomy should be represented as an enlarged taxonomy  $\Psi_{concept}$ , but with considering the semantics of the existing concepts (see Eq. 9):

$$\Psi_{concept} = |\Phi_{\rho} - \Phi_{\alpha}|, \quad (9)$$

where  $\Phi_{\rho}$  includes the concepts detailing the enhanced taxonomy. The enhancement of the number of levels  $\Psi_{level}$  can be done by (see Eq. 10):

<sup>1</sup> <http://www.census.gov/eos/www/napcs>.



**Fig. 4** The hierarchical structure of a taxonomy with dependencies

$$\Psi_{level} = |\Phi_{\rho_{level}} - \Phi_{\alpha_{level}}|, \quad (10)$$

where  $\Phi_{\rho_{level}}$  includes the added levels, and  $\Phi_{\alpha_{level}}$  defines the initial levels.

## Proposed System *TaxoMulti*

The rule-based expert system *TaxoMulti* is customizing taxonomies for channel-specific requirements without misrepresenting the semantic of the underlying domain. First, the method of the underlying system is described and secondly, its implementation including the utilized architecture is presented using this section.

The core advantage of *TaxoMulti* compared to existing works is the usage of different types of mediator concepts operating on different levels. Through this, a highly improved flexibility regarding the adaption process is achieved. A limitation in the adaption process is now only given through the underlying initial taxonomy. Its initial size and fine-granularity define the possible adaption process.

### *TaxoMulti* Method

As taxonomies already use different concept types to express the is-a relationships between concepts, the existing types can be used to define further concept types performing as mediator concepts. In *TaxoMulti*, two types of mediator concepts are used (see Sections “[Taxonomic Dependencies](#)” and “[Taxonomic Collections](#)”) that can be separately asserted with different modification operations (see Section “[TaxoMulti Operations](#)”) on various taxonomic levels. The rule-based expert system allows to combine semantically similar concepts, to split semantically similar concepts, and to enlarge semantically similar concepts. Through this, the mediator concepts are flexible in its appearance and fine-granularity. In addition, as the two novel concept types differ in the existing concept types they are subsuming, all existing concept types are involved. This allows to perform the customization process on taxonomies of any size.

### Taxonomic Dependencies

The concept type *Taxonomic Dependency* relies on the fact, that the sibling concepts sub-ordinating a single concept can have a less generalized semantically sibling relationship between each other than only connecting those with one generalizing super concept (see Fig. 4a, b) [8]. The dependence between sibling concepts can vary from being more likely an antonym (e.g., “Coffee” and “Beer”) or synonym (e.g., “Coffee” and “Tea”), and being a hypernym (“Hot Drink”) or hyponym (“Decaffeinated”) to other sibling concepts (see Fig. 2) [6].

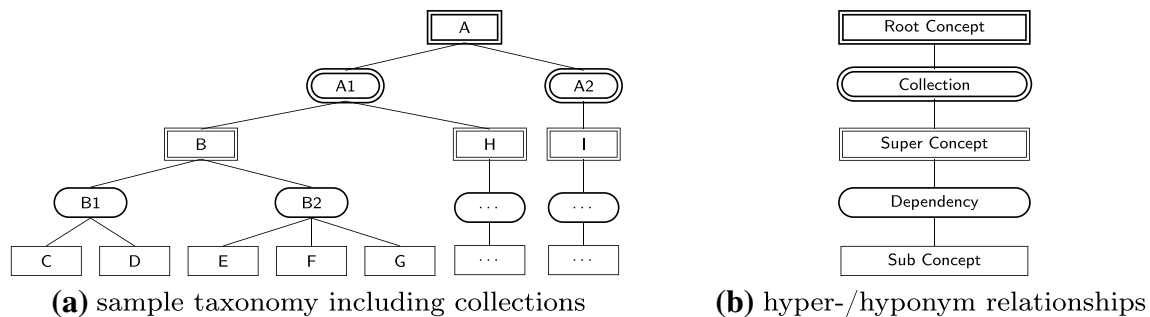
A *Taxonomic Dependency*  $B1$  (shortened as *depof*), is a mediator concept between a super concept  $B$  and a set of sub concepts in  $Y$ , as in Eq. 11, if:

$$B1 = depof(B, Y) : \Leftrightarrow \forall \chi ((\chi \in Y \wedge \chi = subof(\psi)) : \Leftrightarrow \rho > \tau), \quad (11)$$

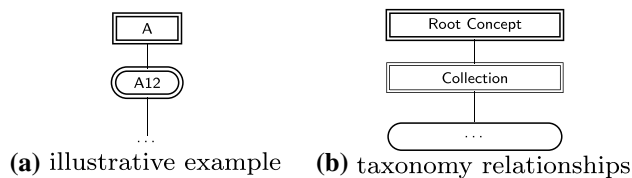
where  $\rho$  is a further semantic relationship between the sibling concepts  $\in Y$ ,  $\psi$  is a super concept, and  $\tau$  is the threshold provided through the provider to verify the relationship between the sibling concepts.

### Taxonomic Collections

The concept type *Taxonomic Collections* focuses on providing more complex relationships between a root concept and its super concepts (see Fig. 5a, b). Through this, a collection concept can include multiple super concepts, depending if the super concepts are semantically similar sibling concepts. For example, the concepts “Seafood” and “Meat Poultry” are more similar compared to the concept “Beverages” (see Fig. 2). Analogues to dependencies, collections offer the possibility to perform a different kind of taxonomic modifications. Consequently, the combination of collections and dependencies ensures that various modification rules can be performed on any concept, and taxonomies of any number of levels can be customized.



**Fig. 5** The hierarchical structure of a taxonomy with collections



**Fig. 6** The modification of the taxonomy with the combine super operation

A *Taxonomic Collection* A1 (shortened as *colof*), is a mediator between a root concept A and a set of super concepts in  $Y$ , as given in Eq. 12, if:

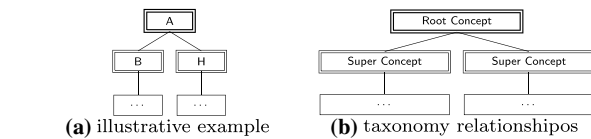
$$A1 = colof(A, Y) : \Leftrightarrow \forall \chi ((\chi \in Y \wedge \chi = subof(v)) : \Leftrightarrow \rho > \tau), \quad (12)$$

where  $\rho$  is a further semantic relationship between the sibling concepts  $\in Y$ ,  $\rho$  is a sub-ordinated concept of the root concept, and  $\tau$  is the threshold provided through the provider to verify the relationship between the sibling concepts.

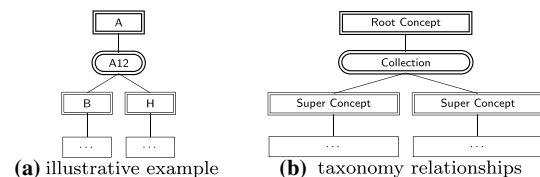
### TaxoMulti Operations

As the mediator concepts are not shown in the initial taxonomy, both above-discussed concept types provide the opportunity to perform different modification operations. Three operations are customizing the initial sub concepts based on the dependency concepts. Another three are used by the super concepts, and are performing on the collection concepts. The rule-based process starts by customizing the super concepts, as those are required to be used as super-ordinated concepts for the resulting sub concepts. Each rule for a dependency can be combined with each rule for a collection, resulting nine possibilities for each taxonomic path to create customized channel-specific taxonomies:

*Combine Super* is reducing the super concepts to the underlying collection, instead of showing the single super concepts (see Fig. 6a, b). This affects a reduction  $\Omega_{R_C}$  of the number of super concepts, shown in Eq. 13:



**Fig. 7** The modification of the taxonomy with the split super operation



**Fig. 8** The modification of the taxonomy with the enlarge super operation

$$\Omega_{R_C} = \Omega_{R_i} - \Omega_{R_N}, \quad (13)$$

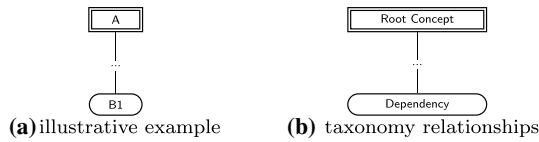
where  $\Omega_{R_C}$  and  $\Omega_{R_i}$  are all concepts being super-ordinated by the same concept,  $\Omega_{R_C}$  are the concepts remaining after the modification is performed,  $\Omega_{R_i}$  is the number of initial concepts super-ordinating a concept, and  $\Omega_{R_N}$  is the number of concepts being combined.

*Split Super* affects that each super concept of a collection is appearing as a single super concept, instead of showing the collection, see Fig. 7a, b. Consequently, this modification affects that the initial taxonomy is shown, meaning a lower reduction as above, but also a lower enlargement as the rule presented as next. Formally, the rule results as given in Eq. 14:

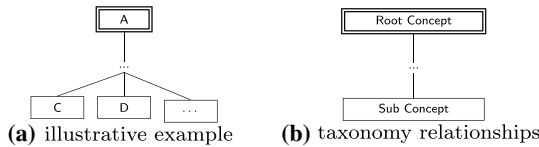
$$\Omega_{R_S} = \Omega_{R_i}, \quad (14)$$

where  $\Omega_{R_S}$  and  $\Omega_{R_i}$  are concepts being super-ordinated by one concept, but  $\Omega_{R_S}$  are the concepts remaining after the modification rule is performed, and  $\Omega_{R_i}$  is the number of initial concepts super-ordinating a concept.





**Fig. 9** The modification of the taxonomy with the combine sub operation



**Fig. 10** The modification of the taxonomy with the splitting sub operation

*Enlarge Super* is using itself to subsume the root concept and all super concepts assigned to it for subsuming the collection, see Fig. 8a, b. It is using the single collection to occur as sub-ordinated concept of the root concept, and the included super concepts to occur below the collection. Correspondingly, it enlarges the number of concepts by the collection, as given in Eq. 15:

$$\Omega_{R_E} = \Omega_{R_i} + \Omega_{R_C}, \quad (15)$$

where  $\Omega_{R_C}$  and  $\Omega_{R_i}$  are concepts being super-ordinated by different concepts resulting in the number of concepts  $\Omega_{R_E}$ . Logically, it increases the number of levels. The number of levels increased is flexible, as it depends on with which other rules it is combined.

*Combine Sub* is summarizing all to the dependency assigned sub concepts, instead of showing the single sub concepts, see Fig. 9a, b. Consequently, this modification affects a reduction  $\Omega_{R_K}$  of the number of sub concepts, as given in Eq. 16:

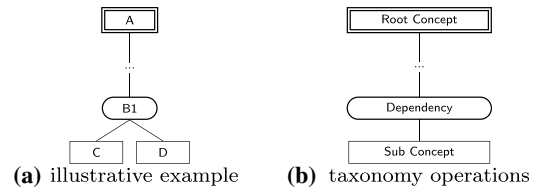
$$\Omega_{R_K} = \Omega_{R_i} - \Omega_{R_N}, \quad (16)$$

where  $\Omega_{R_K}$  and  $\Omega_{R_i}$  are all concepts being super-ordinated by the same concept, but  $\Omega_{R_N}$  are the concepts remaining after the modification rule is performed,  $\Omega_{R_i}$  is the number of initial concepts super-ordinating a concept, and  $\Omega_{R_N}$  is the number of concepts being summarized.

*Split Sub* affects that all sub concepts occur as a single sub concepts, instead of combining the subset to a single dependency, see Fig. 10a, b. Consequently, this modification affects that the initial taxonomy is shown, as given in Eq. 17:

$$\Omega_{R_Z} = \Omega_{R_i}, \quad (17)$$

where  $\Omega_{R_Z}$  and  $\Omega_{R_i}$  are all concepts being super-ordinated by the same concept, but  $\Omega_{R_Z}$  are the concepts remaining after



**Fig. 11** The modification of the taxonomy with the enlarge sub operation

the modification rule is performed, and  $\Omega_{R_i}$  is the number of initial concepts super-ordinating a concept.

*Enlarge Sub* enhances the sub concepts through showing the dependency as additional level, and all included sub concepts, see Fig. 11a, b. It is using the single dependency to occur as sub-ordinated concept of the super concept, and the included sub concepts to occur below the dependency. Correspondingly, it enlarges the number of concepts by the collection, as given in Eq. 18:

$$\Omega_{R_F} = \Omega_{R_i} + \Omega_{R_C}, \quad (18)$$

where  $\Omega_{R_C}$  and  $\Omega_{R_i}$  are concepts being super-ordinate by different concepts resulting in the total number of concepts  $\Omega_{R_F}$ . This increases the number of levels by 1. Along with the rule enlarge super concept, it doubles the number of levels.

## TaxoMulti Implementation

*TaxoMulti* is implemented using logic programming language Prolog. Prolog is known as an effective technique for knowledge engineering and logical reasoning [40]. The reason for this area of operation is that each Prolog program consists of two components [11]:

Knowledge Base	represents information using predicates in the form of facts. For example two facts: the concept “car” is-a (sub concept of) “vehicle”, and the concept “vehicle is-a (sub concept of) “objects”.
Inference Engine	represents predicates in the form of inference rules to infer new knowledge. For example: because “car” is-a (sub concept of) “car”, which is-a (sub concept of) “objects”, “car” is-a (sub concept of) “objects”, too.

In Prolog, the predicates are written in horn clauses. Every predicate consists of a functor, which represents the name of the predicate, and arguments standing in brackets. When using the short form to represent a predicate in Prolog, the number of arguments is written behind the functor (e.g. *concept/2*).

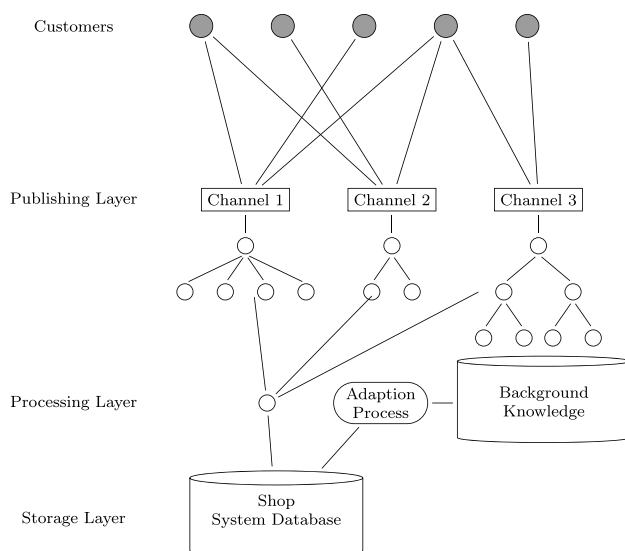


Fig. 12 *TaxoMulti* general architecture consisting of three layers

### Architecture/Layers

The *TaxoMulti* architecture consists of three layers. Only the data displayed on the frontend is customized. This has the benefit that the initial data (in the backend) remains [1]. The layers are (illustrated in Fig. 12):

- **Publishing Layer** is the frontend shown to the multi-channel users. The layer can publish multiple taxonomies, according to the number of channels desired. Of course, the implementation of this layer is not part of the work at hand. This layer is provided by the frontend of the utilized online portal.
- **Processing Layer** forms the actual implementation of *TaxoMulti* and consists of two components. The first component is the knowledge base named *Background Knowledge*. It is required for querying the initial concepts of the *Storage Layer* (see below), and for storing the mediator concepts and semantic relationships between those. The second component named *Adaption Process* is the inference engine. It is used to process different operations for customizing the channel-specific taxonomies. Here, the operator of the multi-channel online portal can define different channels, and the above-mentioned operations as rules, which have to be performed on the mediator concepts.
- **Storage Layer** consists of the database provided by the information management system for storing the initial taxonomy. Logically, the implementation of this layer is not part of the work at hand. This layer is provided by the backend of the utilized online portal. Consequently, our proposed system *TaxoMulti* does not modify the data

stored on this layer. The initial taxonomy is mirrored to the processing layer.

### Background Knowledge

The background knowledge is the first component of the processing layer. It is used for querying the information about the initial taxonomy, storing the information about the mediator concepts and the semantic relationships, as well as about the channels. The initial taxonomy is queried from the storage layer. Each concept regardless of its type (root, super, sub) is mirrored using the same type of fact. This is done to not limit the proposed system to a fixed number of taxonomic levels. The fact to mirror the concepts is named *concept/2* (see Listing 1). It uses its label (*CCL*), along with the unique identifier of a concept (*CCI*). The is-a relationship is expressed using the fact *isa/2*.

#### Listing 1 initial concepts.

```
concept (CCI, CCL).
isa (CCI1, CCI2).
```

The fact *dependency/2* is storing the information about the dependencies using its unique identifier (*DPI*), along with its label (*DPL*) (see Listing 2). A mapping between the dependencies and the super and sub concepts is achieved using the fact *dependencymapping/3*. It uses the unique identifier of the super-ordinated concept (*SPI*), *DPI*, and the unique identifier of the sub-ordinated concept (*SBI*). Analogous to the dependencies, another two facts are used for storing the information about the collections and its mappings. The fact *collection/2* is storing the information about the collections using its unique identifier (*COI*), along with its label (*COL*). A mapping between the collections and the super and sub concepts is achieved using the fact *collectionmapping/3*. It uses the unique identifier of the root concept (*ROI*), *COI*, and the unique identifier of the super concept (*SPI*).

#### Listing 2 mediator concepts and mapping.

```
dependency (DPI, DPL).
dependencymapping (SPI, DPI, SBI).
collection (COI, COL).
collectionmapping (ROI, COI, SPI).
```

### Adaption Process

The adaption process is the second component of the processing layer. It is necessary to finally create channel-specific taxonomies. Doing so, different channels can be defined, and the rules can be applied to the concepts of the initial taxonomy by using the in the previous section explained mediator concepts.



Before finally processing the channel-specific taxonomies, three different facts are needed. The different channels and its operations are stored as fact *channel/2*, where each channel is having a unique identifier (*CNI*), and a name (*CNN*) (see Listing 3). Each channel is having rules to describe the operations to be performed for each collection and dependency: *collectionrule/3*, *dependencyrule/3*. Each fact includes the identifier of the channel, of the mediator concept, and an identifier for the desired operation (*OPP*, or *OPB*).

**Listing 3** channels and rules.

```
channel(CNI, CNN).
rulesuperconcept(CNI, COI, OPP).
rulesubconcepts(CNI, DPI, OPB).
```

The final processing of the channel-specific taxonomies is done in four steps (see Listing 4). It starts with asking the user for the channel (*user\_input(CNI)*) to be modified. Afterwards, the rule *isa\_operation* is performed, which creates maximum paths in the form *isa(ROI, COI, SPI, DPI, SBI)*. Based on those, the final modifications are performed using the rule *modification(CNI)*. Depending on the rule for a specific mediator concept, the paths are modified accordingly.

**Listing 4** s/0

```
dynamicadaption:-
    user_input(CNI), isa_operation,
    modification(CNI), removeduplicates(CNI).

modification(CNI):-
    channel(CNI, _),
    forall(isa(ROI, COI, SPI, DPI, SBI),
        (
            comsupcomsub(CNI, ROI, COI, SPI, DPI, SBI);
            comsupsplsub(CNI, ROI, COI, SPI, DPI, SBI);
            comsupenlsub(CNI, ROI, COI, SPI, DPI, SBI);
            splsupcomsub(CNI, ROI, COI, SPI, DPI, SBI);
            splsupsplsub(CNI, ROI, COI, SPI, DPI, SBI);
            splsupenlsub(CNI, ROI, COI, SPI, DPI, SBI);
            enlsupcomsub(CNI, ROI, COI, SPI, DPI, SBI);
            enlsupsplsub(CNI, ROI, COI, SPI, DPI, SBI);
            enlsupenlsub(CNI, ROI, COI, SPI, DPI, SBI))).
```

## Case Study and Demonstration

As an illustration of the proposed rule-based expert system *TaxoMulti*, let us consider the retailing market *Northwind*.<sup>2</sup> The retailing market operates a multi-channel e-commerce online portal. In total, three digital channels with the accompanying device classes are distinguished according to the underlying screen size (summarized in Table 2):

**Table 2** Channels used by the retailing market *Northwind*

Channelid	Name
1	Tablet
2	Mobile
3	Desktop

- *Mobile* is the channel used by most customers for quick browsing. However, many customers are terminate the browsing. The goal is to reduce the taxonomy published on this channel to allow a more comfortable browsing experience.
- *Tablet* is the channel recently used by many customers for placing orders. The goal is that this channel is using the initial taxonomy. No change should be made as the channel operates successful.
- *Desktop* is the channel used by some customers. Customers looking for specific things are mainly using this channel. The goal is that this channel is using the maximum taxonomy to allow a detailed navigation for all customers.

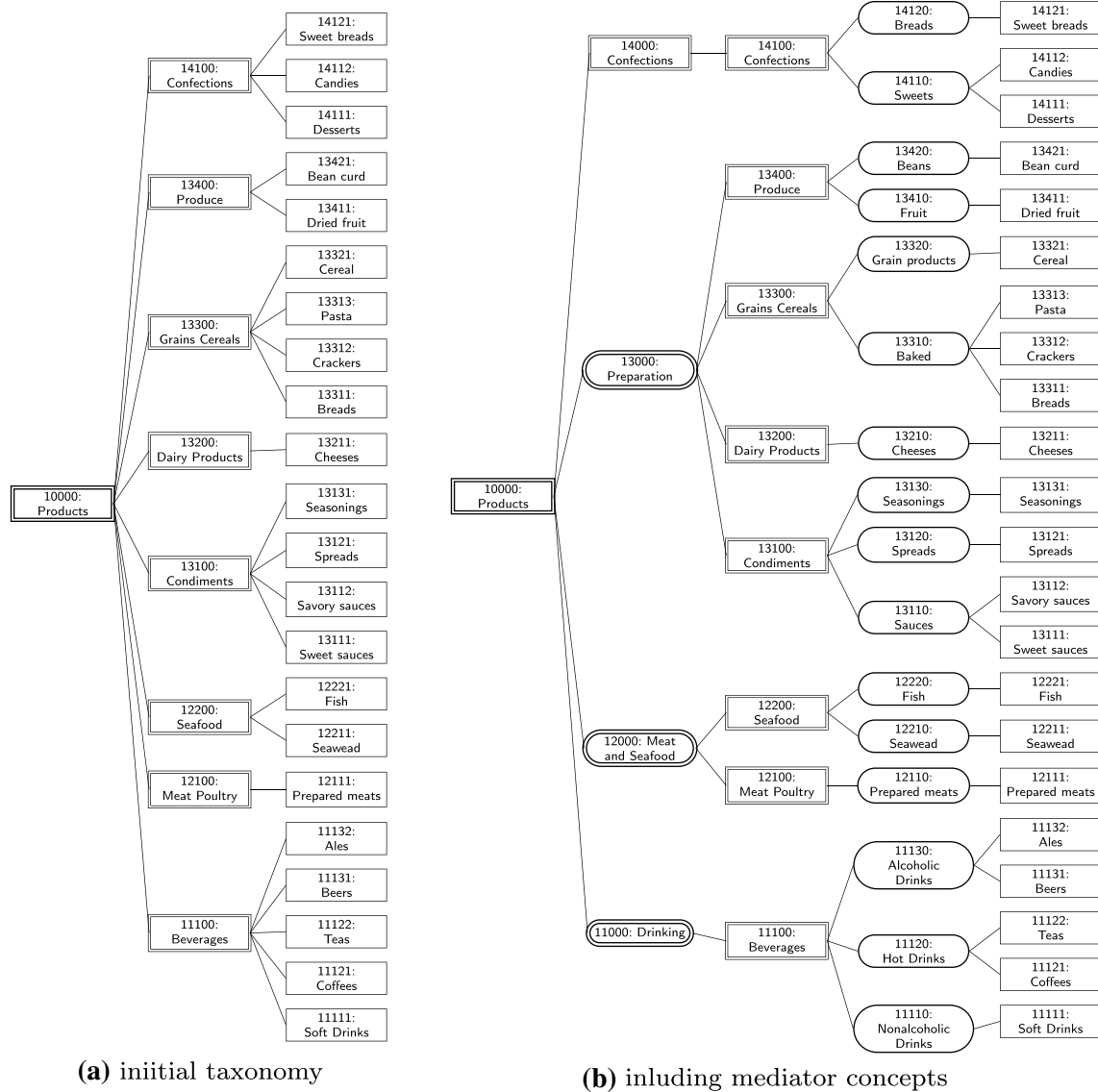
Based on the initial taxonomy, mediator concepts are created (see Fig. 13).

## Mobile Channel-Specific Taxonomy

As the complete range of concepts is too large to be shown on a mobile device, the taxonomy published in the frontend should be reduced. To do so, mediator concepts performing on the different levels (super and sub concepts) are assigned with corresponding rules. The collections performing on the super concepts are assigned with the rule combining similar super concepts, see Table 3. The dependencies performing on the level of sub concepts are assigned with the rule combining similar sub concepts, see Table 4.

Based on the initial taxonomy, in total eight super concepts exist. Depending on the semantic similarity between the super concepts, a different number of super concepts can sub-ordinate the different collections. The super concept “Confections” for example and “Beverages” have no similar super concepts. In contrast, the super concept “Produce” for example sub-ordinates the new collection “Preperation”, as well as its sibling concepts “Grains Cereals”, “Dairy Products”, and “Condiments”. Same for the super concept “Seafood”. Along with its sibling concept “Meat Poultry” it is classified into the collection “Meat and Seafood”. The new labels for the collections indicate that it has a less semantic meaning compared to the super concepts. For example, the label “Meat and Seafood” is more general than to split the concept along with its label into “Seafood” and “Meat Poultry”. Correspondingly,

<sup>2</sup> <https://northwinddatabase.codeplex.com/>.



**Fig. 13** The taxonomy of *Northwind* including mediator concepts

**Table 3** Channel-specific rules on the level of super concepts for the mobile taxonomy

Channelid	Collectionid	Operation
2	30001	Combine Super
2	30002	Combine Super
2	30003	Combine Super
2	30004	Combine Super

having different semantic weights for one level would not be meaningful, as it hampers the interpretation of the domain by the user. For this reasons, the super concepts having no similar sibling concept are also categorized into a collection carrying a more general label. For this

reason, “Confections” sub-ordinates the collection “Sweet food”, and “Beverages” sub-ordinates the collection “Drinking”. In total, the number of super concepts can be halved ( $Y_{super} = 8 - 4$ ). On the level of the sub concepts, in total 22 less generalized concepts exist. Again, a different number of sub concepts can sub-ordinate the different dependencies, due to the semantic similarity between the sub concepts. Hereby, the number of sub concepts sub-ordinate a dependency goes from one to three. In total, the number of sub concepts can be reduced by six concepts ( $Y_{sub} = 22 - 16$ ). In the end, the complete number of concepts (except the root concept) can be reduced by ten concepts ( $Y = 30 - 20$ ), which indicates a relative minimum reduction by 33.33 %, see Fig. 14.

**Table 4** Channel-specific rules on the level of sub concepts for the mobile taxonomy

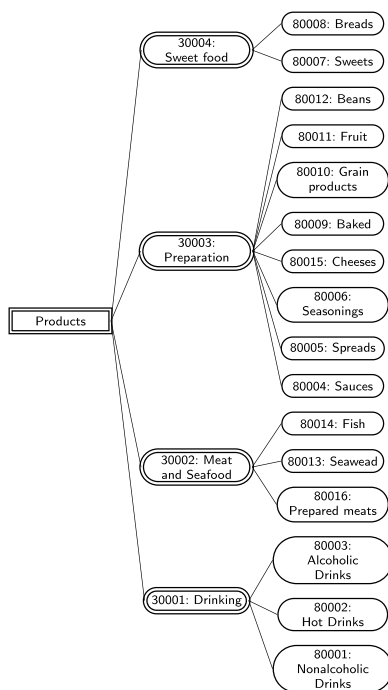
Channelid	Dependencyid	Operation
2	80001	Combine Sub
2	80002	Combine Sub
2	80003	Combine Sub
2	80004	Combine Sub
2	80005	Combine Sub
2	80006	Combine Sub
2	80007	Combine Sub
2	80008	Combine Sub
2	80009	Combine Sub
2	80010	Combine Sub
2	80011	Combine Sub
2	80012	Combine Sub
2	80013	Combine Sub
2	80014	Combine Sub
2	80015	Combine Sub

**Table 5** Channel-specific rules on the level of super concepts for the printed taxonomy

Channelid	Collectionid	Operation
2	30001	Split Super
2	30002	Split Super
2	30003	Split Super
2	30004	Split Super

**Table 6** Channel-specific rules on the level of sub concepts for the printed taxonomy

Channelid	Dependencyid	Operation
2	80001	Split Sub
2	80002	Split Sub
2	80003	Split Sub
2	80004	Split Sub
2	80005	Split Sub
2	80006	Split Sub
2	80007	Split Sub
2	80008	Split Sub
2	80009	Split Sub
2	80010	Split Sub
2	80011	Split Sub
2	80012	Split Sub
2	80013	Split Sub
2	80014	Split Sub
2	80015	Split Sub

**Fig. 14** The reduced taxonomy used for the channel “mobile”

## Desktop Channel-Specific Taxonomy

As nowadays, desktop computers usually have extra-large screens, users using this channel should see the most enlarged taxonomy to better filter for the concepts they require. As the goal is to achieve the maximum size of the taxonomy, it is the one including all of our defined mediator concepts (see Fig. 13b). For this reason, the modification

rules to enlarge the taxonomy should be enlarged for all levels. The collections performing on the level of super concepts are assigned with the rule to enlarge the super concepts, see Table 5. The dependencies performing on the level of sub concepts are assigned with the rule enlarging similar sub concepts, as given in Table 6.

As explained in the previous section, in total eight super concepts exist. Such super concepts are assigned to the collections as described above, namely to four collections. The difference to the mobile taxonomy is that now the super concepts are not combined. The super concepts are displayed as single concepts, in total eight. In addition, the four collections are also displayed to the user below the root concept. Similar is the modification for the sub concepts. In contrast to the mobile taxonomy, the sub concepts are not reduced to dependency. The dependencies are used as an additional level between the super and sub concepts. Correspondingly, the sub concepts remain as single sub concepts. In the end, the complete enhanced taxonomy includes 20 more concepts ( $\Psi = 50 - 30$ ) and two additional levels ( $\Psi_{level} = 5 - 3$ ). Through this, the complete relative enhancement is 66.67 % with respect to the number of concepts. And 66.67 % of

enlargement was achieved with respect to the number of levels.

## Experimental Evaluation

The evaluation of rule-based attempts is often considered as challenging and differs significantly compared to other disciplines [33, 41]. Often, the standard quality metrics (e.g. accuracy, f-measure) are not in the focus because of different reasons [32]. First, the rule-based systems are implemented to support or automate a single task. So, the rules to be included heavily depend on the human expert. Second, the implementation consists not only of an algorithm, but of a knowledge base and an inference engine forming the background knowledge [11]. So, a mutual supplementation of both components is decisive. However, both components significantly depend on the domain and the experts opinion, which is not qualitatively measurable. Because of the mentioned reasons, the evaluation of the proposed rule-based expert system *TaxoMulti* is multi-faceted and intense. In detail, the following metrics are used:

- *Reduction Efficiency* is measured to verify the reduction of the taxonomy for channels requiring a reduced taxonomy. Hereby, the system is evaluated against the provided problem formulation (see Sect. 2.2). To do so, each taxonomy is firstly investigated against its initial number of sub ( $\Phi_{\alpha_{sub}}$ ) and super ( $\Phi_{\alpha_{super}}$ ) concepts. Secondly, the number of sub ( $\Phi_{\beta_{sub}}$ ) and super ( $\Phi_{\beta_{super}}$ ) concepts for the reduced taxonomy are measured. This results in the reduction of sub ( $Y_{sub}$ ) and super ( $Y_{super}$ ) concepts, and finally the reduction of the complete taxonomy ( $Y$ ), respectively. The reduction efficiency is in addition compared with the latest work on personalized directories, presented by [6], in the following named Personalized Directories. We do this by comparing their proposed operations to reduce the taxonomy, with the operations proposed in the work at hand.
- *Enlargement Efficiency* is computed to verify the enlargement of the taxonomy for channels requiring a larger taxonomy. We investigate the system against the provided problem formulation (see Sect. 2.2). First, the initial number of super and sub concepts is measured ( $\Phi_{\alpha}$ ), which is compared with the enhanced number of concepts ( $\Phi_{\rho}$ ) to result in the enhancement with respect to the number of concepts ( $\Psi_{concept}$ ). In addition, the enhancement with respect to the number of levels is computed ( $\Psi_{level}$ ). To do so, the enhanced number of levels ( $\Phi_{\rho_{level}}$ ) is compared with the initial number of levels ( $\Phi_{\alpha_{level}}$ ). Both final metrics are compared against the most recent operations presented in Personalized Directories.

- *Analytical Comparison* of the rule-based system is performed with the help of the related research paradigms aiming to modify taxonomies: dynamic taxonomies, catalog segmentation, and personalized directories. The work at hand and the mentioned paradigms are investigated according to four criteria: if a taxonomy reduction can be performed by the techniques if the reduction is performed without changing the domain, if the taxonomy enhancement can be performed, and if the enlargement is performed without changing the domain.
- *Quantitative Comparison* is performed based on the analysis above, for all techniques being capable to reduce the taxonomy, respectively to enlarge the taxonomy. This is done by firstly analyzing the minimal possible relative reduction ( $Y_{min}$ ), with (see Eq. 19):

$$Y_{min} = \left\{ 1 - \left( \frac{\Phi_{\beta_{min}}}{\Phi_{\alpha}} \right) \right\} \times 100, \quad (19)$$

where  $\Phi_{\beta_{min}}$  is the remaining number of concepts if a minimal subset of the taxonomy is reduced, respectively enlarged. Correspondingly, the maximum possible relative reduction and enlargement ( $Y_{max}$ ) is also measured, with (see Eq. 20):

$$Y_{max} = \left\{ 1 - \left( \frac{\Phi_{\beta_{max}}}{\Phi_{\alpha}} \right) \right\} \times 100, \quad (20)$$

where  $\Phi_{\beta_{max}}$  is the remaining number of concepts the maximum reduction or enlargement capability is performed. For both,  $Y_{min}$  and  $Y_{max}$ , the remaining domain is investigated. Namely, if the performed reduction/enlargement is changing the semantics of the domain ( $\Xi$ ), with (see Eqs. 21 and 22):

$$\Xi_{min} = \left\{ 1 - \left( \frac{\Theta_{min}}{\Phi_{\alpha}} \right) \right\} \times 100, \quad (21)$$

$$\Xi_{max} = \left\{ 1 - \left( \frac{\Theta_{max}}{\Phi_{\alpha}} \right) \right\} \times 100, \quad (22)$$

where  $\Xi_{min}$  is the relative remaining domain if the minimal reduction/enlargement is performed,  $\Xi_{min}$  is the remaining domain if the maximal possible reduction/enlargement is performed, and  $\Phi_{\alpha}$  is the complete domain. Finally, both metrics,  $Y$  and  $\Xi$  are multiplied resulting  $\Omega$ , the efficiency of the technique with respect to the remaining domain, with (see Eq. 23):

$$\Omega_R = |\Omega_{min} - \Omega_{max}| = |Y_{min} \times \Xi_{min} - Y_{max} \times \Xi_{max}|, \quad (23)$$

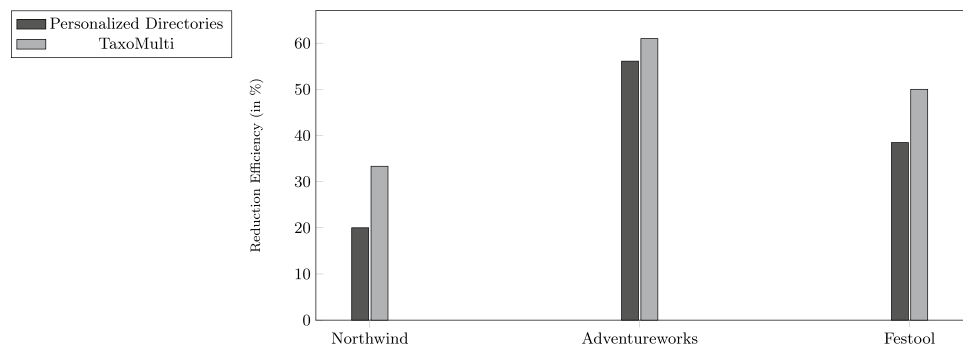
where  $\Omega_{min}$  is the minimum efficiency, and  $\Omega_{max}$  is the maximum efficiency. Consequently, the higher  $\Omega$  (0 to 1), the more capable the technique is to be used for

**Table 7** Characteristics and parameters of the databases used for experimental results

Concept type	Adventure works	Northwind	Festool
Super concepts	4	8	9
Sub concepts	37	22	43
Dependencies	14	16	23
Collections	2	4	3

**Table 8** Taxonomy reduction efficiency of *TaxoMulti*

Variable	Northwind		AdventureWorks		Festool	
	$\bar{a}^a$	$r^b$	$\bar{x}^a$	$r^b$	$\bar{a}^a$	$r^b$
$Y$	10	33.33	25	60.98	26	50
$\Phi_\alpha$	30	100	41	100	52	100
$\Phi_\beta$	20	66.67	16	39.02	26	50
$Y_{sub}$	6	27.27	23	62.16	20	46.51
$\Phi_{\alpha_{sub}}$	22	100	37	100	43	100
$\Phi_{\beta_{sub}}$	16	72.73	14	37.84	23	53.49
$Y_{super}$	4	50	2	50	6	66.67
$\Phi_{\alpha_{super}}$	8	100	4	100	9	100
$\Phi_{\beta_{super}}$	4	50	2	50	3	33.33

<sup>a</sup> Absolute Value<sup>b</sup> Relative Value in %**Fig. 15** Reduction efficiency of *TaxoMulti* for three different database

omni-channel retailing. Finally, a comparison between  $\Omega_{min}$  and  $\Omega_{max}$  is resulting the range  $\Omega_R$  of possible reduction/enlargement, to show the flexibility of the technique.

To ensure that all the above-mentioned analysis are investigated in different directions, the proposed rule-based expert system is applied to different databases. In detail, two open (*AdventureWorks*,<sup>3</sup> *Northwind*), and one database provided by a German company (*Festool*<sup>4</sup>) are used. The characteristics of the databases are summarized in Table 7.

<sup>3</sup> <http://msftdbprodsamples.codeplex.com/>.

<sup>4</sup> <https://www.festool.de/>.

## Reduction Efficiency

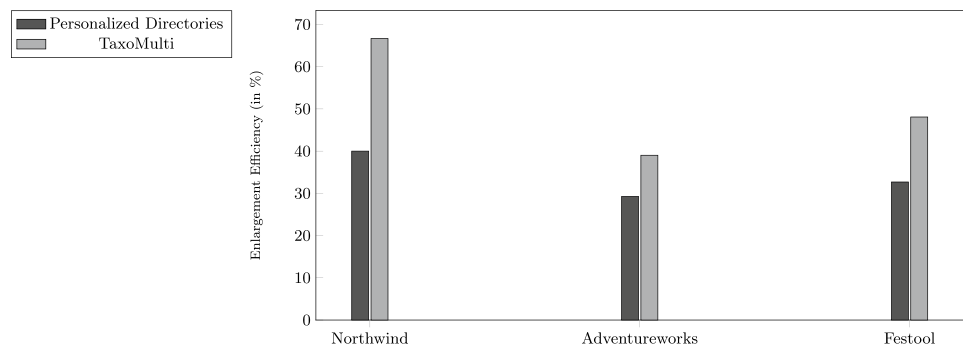
When summarizing the results for the taxonomy reduction, the taxonomies were reduced with the proposed system by on average 50%, see Table 8.

The results obtained reveal that the possible reduction does not depend on the general size of the taxonomy. The efficiency mainly depends on the fine-granularity of the taxonomy for the different levels of the initial taxonomy. The more similar the concepts for a single super-ordinate concept

are, the higher the taxonomy can be reduced. Through this, there is no linear correlation between the size of the initial taxonomy and the reduction. Small, as well as very large taxonomies can be reduced with the same high efficiency. A higher efficiency can be further achieved when using more general mediator concepts, which is also possible with the proposed system.

Compared to the taxonomy operations presented in Personalized Directories, a further decrease of 9.93% is performed, see Fig. 15. This is achieved through using the novel concept type collection, which is acting as a mediator concept between the root and a set of super concepts. The works in Personalized Directories lack a mediator concept acting



**Fig. 16** Enlargement efficiency of *TaxoMulti* for three different databases**Table 9** Taxonomy enlargement efficiency of *TaxoMulti*

Variable	Northwind		AdventureWorks		Festool	
	$\bar{a}^a$	$r^b$	$\bar{x}^a$	$r^b$	$\bar{a}^a$	$r^b$
$\Psi_{concept}$	20	66.67	16	39.02	23	48.08
$\Phi_{\rho}$	30	100	41	100	52	100
$\Phi_{\alpha}$	50	166.67	57	139.02	77	148.08
$\Psi_{level}$	2	66.67	2	66.67	2	66.67
$\Phi_{\rho_{level}}$	3	100	3	100	3	100
$\Phi_{\alpha_{level}}$	5	166.67	5	166.67	5	166.67

<sup>a</sup> Absolute Value<sup>b</sup> Relative Value in %

between the root and super concepts. In the case of requiring a very reduced taxonomy as in our experiment, the different sets of super concepts are reduced to the underlying collections. A further reduction or a lower reduction could have only been achieved if the initial taxonomy is already bigger or smaller so that the size of the sets change completely. However, manipulating the initial taxonomy is not the scope of this work, as the taxonomy is often used in other entrepreneurial perspectives for other tasks as well (e.g. product master data) [10, 61].

## Taxonomy Enlargement

When summarizing the results corresponding to the taxonomy enlargement, an efficiency greater 50% was achieved for the number of concepts, and an efficiency of almost 67% for the number of levels, see Table 9. The resulted value for enlarging the number of concepts is analogous to the subtraction of the reduction, which shows the simplicity of the proposed operations, as the mediator concepts can be used in multiple directions.

Compared to the operations presented in Personalized Directories, an increase of 17.29% was measured with respect to the number of concepts, see Fig. 16. This is achieved through allowing the flexible combination of two types of mediator concepts, whereas the works on Personalized Directories only adapt each concept type separately. In the case of requiring a very large taxonomy as in our

experiment, the different sets of super concepts that are bound to one collection are shown, including its super-ordinated concept. Additionally, the different sets of sub concepts that are bound to one dependency are shown, including its super-ordinated concept. A further increase or a lower increase could have only been achieved if the initial taxonomy is already bigger or smaller, so that the size of the sets change. However, as discussed, changing the initial taxonomy is not expedient.

## Analytical Comparison

For demonstrating the strength and weakness of *TaxoMulti* regarding the adaption of taxonomies, the systems main criteria are compared against other research paradigms aiming to adapt taxonomies: dynamic taxonomies, catalog segmentation, and personalized directories. To give the reader the most comprehensive comparison, we first explain the different paradigms in more detail, before identifying the strength and weakness of *TaxoMulti*.

### Analytical Techniques Paradigms

A dynamic taxonomy prunes itself in response to the request and so considers the significance of a user-query [50]. These paradigm has been proposed as a solution to combine navigation and querying, offering both

**Table 10** Comparison with most related work in related paradigms

Paradigm (recent work)	Taxonomy reduction		Taxonomy enlargement	
	Capability	Semantic	Capability	Semantic
Dynamic Taxonomies [29]	✓			
Catalog Segmentation [52]	✓			
Personalized Directories [6]	✓	✓	✓	✓
<i>TaxoMulti</i>	✓	✓	✓	✓

expressivity and interactivity, e.g. in [15, 24, 29, 31, 50, 56, 58]. However, as the (semi)-automatic technique is until now always based on the initial taxonomy, this paradigm is not capable to enlarge the taxonomy. In addition, as not detected matches against the provided keyword are not displayed inside the dynamic taxonomy, the domain is reduced accordingly.

Catalog segmentation is also based on a (static) taxonomy, but proposes to create a variety of different sub-taxonomies for different segments of users, e.g. in [4, 36, 37, 39, 52, 63]. Its capabilities lie on a high effective reduction of the taxonomy, but the computed sub-taxonomies are losing information about the domain. Logically, enlarged taxonomies are not considered in this paradigm.

Personalized directories are using different taxonomy modification rules, i. a. operations to represent the taxonomy according to a user-specific requirement, e.g. in [6, 21, 22, 28, 34, 47, 62]. Hereby, the main focus is to effectively reduce the distance of preferred concepts to the root distance. This paradigm does not lose information about the domain, and the enlargement would be possible if the in this work presented mediator concepts are used.

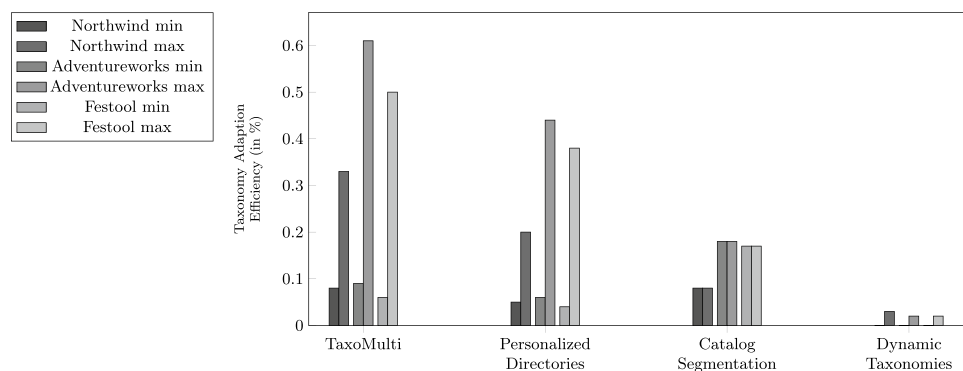
## Analytical Strength and Weakness Comparison

The proposed rule-based system could outperform all existing paradigms according to the studied criteria, see Table 10, whereby the paradigm of personalized directories is closest to our proposed system. However, the above-provided comparison of the related works offers to identify the strength of the system at hand, but also its possible further extensions:

- None of the other recent techniques provides an effective solution to adapt taxonomies for supporting taxonomies in omni-channel. The main difference to other works relies on the usage of the mediator concepts. Those are flexible in its appearance, and can be used to reduce the taxonomy, as well as to enlarge the taxonomy. Hereby, both different cases can be achieved without missing semantically information about the domain.
- The other related paradigms offer techniques to adapt the taxonomy according to user feedback. The most recent work on personalized directories includes a recommender system to automatically perform the taxonomical operations. In *TaxoMulti*, the operation is automatically performed, but the status for the mediator concepts to perform the rule has to be provided by the expert.

## Quantitative Comparison

Based on the above studied analytical comparison, a quantitative comparison can be performed for the techniques being capable to reduce and/or enlarge the taxonomy. The techniques being capable to reduce the taxonomy are: *TaxoMulti*, Personalized Directories, Catalog Segmentation, and Dynamic Taxonomies, see Fig. 17. The techniques being capable to enlarge the taxonomy are: *TaxoMulti*, and Personalized Directories. However, a comprehensive comparison of these both techniques is already presented above. The

**Fig. 17** Taxonomy adaption efficiency comparison result regarding domain change

quantitative comparison to reduce the taxonomy with respect to the reduction of the domain allows three observations:

1. *TaxoMulti* shows overall the best results for three different databases regarding the maximum reduction of the taxonomy. The main reason for the improvement compared to catalog segmentation and dynamic taxonomies is that *TaxoMulti* is not reducing the taxonomy semantically, but numerically. Compared to personalized directories, the improvement was performed because of *TaxoMulti* is capable to reduce the taxonomy on all existing levels. For all three databases, *TaxoMulti* resulted a minimum reduction of the taxonomy with respect to the remaining of the domain, of on average 48%, the personalized directories technique of 38%, the catalog segmentation technique of 14%, and the technique of dynamic taxonomies of 2%. So compared to the currently best performing existing works, *TaxoMulti* has achieved an improvement of + 26,31 %.
2. Only the technique of catalog segmentation shows better results compared to *TaxoMulti* regarding the minimal reduction of the taxonomy. This is, because it is assumed, that the minimum reduction for catalog segmentation techniques is performed for a minimum one super concept including the sub concepts, but *TaxoMulti* still shows the semantics of all super concepts, including the sub concepts. Compared to the latest techniques on personalized directories, a further improvement was performed using *TaxoMulti* instead. The improvement is again affected, because of *TaxoMulti* can reduce the taxonomy for each level. For all three databases, *TaxoMulti* resulted in a maximum reduction of the taxonomy with respect to the remaining of the domain, of on average 8%, the personalized directories technique of 5%, the catalog segmentation technique of 18%, and dynamic taxonomies of 0%. Compared to currently best performing existing works, *TaxoMulti* has achieved an improvement of + 60,00 %.
3. *TaxoMulti* shows the best performance regarding flexibility. Catalog segmentation fully lacks with respect to flexibility. The reason is, the more the number of concepts of the taxonomy is reduced, the more, the domain is also reduced semantically. Similar for dynamic taxonomies, where the techniques are reducing the taxonomy regarding the precision of a provided keyword. However, the more precise the keyword is, the more the taxonomy is also reduced semantically. In contrast, the more vague the keyword is, the bigger the semantics of the taxonomy is, but without any reduction. For all three databases, on average, *TaxoMulti* resulted in a range of 40%, the techniques of personalized directories a range of 33%, the techniques of catalog segmentation a range of 4%, and the technique of dynamic taxonomies, a range of 2%.

So compared to the currently best performing existing works, *TaxoMulti* has achieved an improvement of + 21,21 %.

## Conclusions

This work has presented *TaxoMulti*, the first rule-based expert system to customize product taxonomies for multi-channel e-commerce. Contrary to previous work on catalog segmentation, dynamic taxonomies, or personalized directories, *TaxoMulti* is not reducing the taxonomy in a semantically manner. This progress is achieved by using flexible mediator concepts.

Through the adaption of the taxonomy according to the channels' specification, the rule-based system remedies three major drawbacks the formal taxonomies suffer from. Firstly and foremost, the system provides an environment to flexible customize taxonomies. This is achieved through providing various and combinable taxonomic operations, which perform on the different levels of the taxonomy. The flexibility of these operations is achieved by performing on mediator concepts. Those are flexible in its appearance and can be asserted with channel-specific rules. The efficiency of the mediator concepts has been evident by an evaluation performed on three databases. For channels requiring a smaller taxonomy (e.g. mobile), the taxonomy can be reduced by on average almost 50%. For channels requiring a larger taxonomy (e.g. desktop), the taxonomy can be enlarged by over 50% in the number of concepts. In addition, an enlargement for the number of levels of on average almost 67% was achieved. Compared to the currently best performing existing works, *TaxoMulti* has achieved an improvement of + 26,31 % for the reduction of the taxonomy, + 60 % for the enlargement of the taxonomy, and + 21,21 % in terms of flexibility.

Future work on *TaxoMulti* can be divided in two directions. By applying recommender system techniques, the taxonomy can be adapted automatically according to the channel and users' preferences. And, by combining the adapted taxonomy with user interface design, so-called mixed-reality scenarios could be investigated, e.g. in [38].

**Funding** Open Access funding enabled and organized by Projekt DEAL.

## Declarations

**Conflict of interest** The authors declare that they have no conflict of interest.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- van der Aalst W, Bichler M, Heinzl A. Robotic process automation. *Bus Inf Syst Eng*. 2018;60:269–72.
- Abbas A, Bilal K, Zhang L, Khan S. A cloud based health insurance plan recommendation system: a user centered approach. *Fut Gen Comput Syst*. 2015;43–44:99–109.
- Ali F, El-Sappagh S, Riazul Islam S, Ali A, Attique M, Imran M, Kyung-Sup K. An intelligent healthcare monitoring framework using wearable sensors and social networking data. *Fut Gen Comput Syst*. 2021;114:23–43.
- Amiri A. Customer-oriented catalog segmentation: effective solution approaches. *Decis Support Syst*. 2006;42:1860–71.
- Angermann H. Manager's guide to sharepoint server 2016. New York: Springer; 2017.
- Angermann H, Ramzan N. Taxopublish: towards a solution to automatically personalize taxonomies in e-catalogs. *Expert Syst Appl*. 2016;66:76–94.
- Angermann H, Ramzan N. Taxonomy matching using background knowledge - linked data, semantic web, and heterogeneous repositories. New York: Springer; 2017.
- Angermann H, Pervez Z, Ramzan N. Taxo-semantics: assessing similarity between multi-word expressions for extending e-catalogs. *Decis Support Syst*. 2017;98:10–25.
- Beck N, Rygl D. Categorization of multiple channel retailing in multi-, cross-, and omni-channel retailing for retailers and retailing. *J Retail Consum Serv*. 2015;27:170–8.
- Bovea M, Perez-Belis V. A taxonomy of ecodesign tools for integrating environmental requirements into the product design process. *J Clean Prod*. 2012;20(1):61–71.
- Bramer M. Logic programming with Prolog. New York: Springer; 2014.
- Brown B, Bleecker J, D'adamo M, Ferreira P, Formo J, Gloss M, Holm M, Hook K, Johnson E, Kaburuan E, et al. (2016) The ikea catalogue: Design fiction in academic and industrial collaborations. In: Proceedings of the 19th International Conference on Supporting Group Work, pp 335–344
- Brown J, Dant R (2014) The role of e-commerce in multi-channel marketing strategy. In: Handbook of strategic e-business management, Springer, pp 467–487
- Cai Y, Lo C. Omni-channel management in the new retailing era: a systematic review and future research agenda. *Int J Prod Econ*. 2020;229.
- Calegari S, Pasi G. Personal ontologies: generation of user profiles based on the YAGO ontology. *Inf Process Manag*. 2013;49:640–58.
- Cao L, Li L. The impact of cross-channel integration on retailers' sales growth. *J Retail*. 2015;91:198–216.
- Chen D, Li X, Zhang J (2010) User-oriented intelligent service of e-catalog based on semantic web. In: Proceedings of the 2nd IEEE International Conference on Information Management and Engineering, IEEE, Chengdu, China, pp 449–453
- Cverdelj-Fogarasi I, Sladic G, Gostojic S, Segedinac M, Milosavljevic B. Semantic integration of enterprise information systems using meta-metadata ontology. *IseB*. 2017;15:1–48.
- Euzenat J, Shvaiko P. Ontology matching. 2nd ed. New York: Springer; 2013.
- Farghaly K, Abanda F, Vidalakis C, Wood G. Taxonomy for bim and asset management semantic interoperability. *J Manag Eng*. 2018;34:04018012-1–13.
- Farsani H, Nematbakhsh A. Designing a catalog management system - an ontology approach. *Malaysian Journal of Computer Science (Faculty of Computer Science and Information Technology)*. 2007;20:119–27.
- Fathy N, Badr N, Gharib T (2014) Personalized web search ranking based on different information resources. In: Signal processing and information technology, Springer, pp 211–214
- Fellbaum C. WordNet: an electronic lexical database. Cambridge, MA: MIT-Press; 1998.
- Gollub T, Voelske M, Hagen M, Stein B (2014) Dynamic taxonomy composition via keyqueries. In: Proceedings of the 14th ACM/IEEE-CS Joint Conference on Digital Libraries (ACM/IEEE), pp 39–48
- Gruninger M, Bodenreider O, Olken F, Obrst L, Yim P. Ontology summit 2007-ontology, taxonomy, folksonomy: understanding the distinctions. *Appl Ontol*. 2008;3(3):191–200.
- Herhausen D, Kleinlercher K, Verhoef P, Emrich O, Rudolph T. Loyalty formation for different customer journey segments. *J Retail*. 2019;95:9–29.
- Hosseini S, Merz M, Roeglinger M, Wenninger A. Mindfully going omni-channel: an economic decision model for evaluating omni-channel strategies. *Decis Support Syst*. 2018;109:74–88.
- Joh Y, Lee J. Buyer's customized directory management over seller's e-catalogs: logic programming approach. *Decis Support Syst (Elsevier)*. 2003;34:197–212.
- Johnston V, Black M, Wallace J, Mulvenna M, Bond R (2019) A framework for the development of a dynamic adaptive intelligent user interface to enhance the user experience. In: Proceedings of the 31st European Conference on Cognitive Ergonomics (ACM), pp 32–35
- Kim J, Chun S. Cannibalization and competition effects on a manufacturer's retail channel strategies: implications on an omni-channel business model. *Decis Support Syst*. 2018;109:5–14.
- Kumaraguru S, Rachuri S, Lechevalier D. Faceted classification of manufacturing processes for sustainability performance evaluation. *Int J Adv Manuf Technol (Springer)*. 2014;75:1309–20.
- Lavrac N, Flach P, Zupan B (1999) Rule evaluation measures: a unifying view. In: Inductive Logic Programming, Springer, New York, pp 174–185
- Ligeza A. Logical foundations for rule-based systems. 2nd ed. New York: Springer; 2006.
- Lin C, Hong C. Using customer knowledge in designing electronic catalog. *Expert Syst Appl*. 2008;34:119–27.
- Luan Y, Zhang Z. Research on e-commerce integrated management information system of cross-border enterprises based on collaborative information middleware. *IseB*. 2020;18:527–43.
- Mahdavi I, Movahednejad M, Adbesh F. Designing customer-oriented catalogs in e-CRM using an effective self-adaptive genetic algorithm. *Expert Syst Appl*. 2011;38:631–9.
- Mark T, Bulla J, Niraj R, Bulla I, Schwarzwaeiler W. Catalogue as a tool for reinforcing habits: empirical evidence from a multichannel retailer. *Int J Res Mark*. 2019;36:528–41.
- Meegahapola L, Perera I (2017) Enhanced in-store shopping experience through smart phone based mixed reality application. In: 2017 Seventeenth International Conference on Advances in ICT for Emerging Regions, pp 1–8

39. Melo T, Silva A, Moura E, Calado P. Opinionlink: leveraging user opinions for product catalog enrichment. *Inf Process Manag.* 2019;56:823–43.
40. Merritt D. Building expert systems in Prolog. Heidelberg: Springer; 1989.
41. Miranda P, Isaias P, Crisostomo M (2007) Expert systems evaluation proposal. In: Symposium on Human Interface and the Management of Information, Springer, New York, pp 98–106
42. Müller-Lankenau C, Wehmeyer K, Klein S (2005) Strategic channel alignment - perspectives on the combination of physical and virtual distribution channels. In: Information Systems and e-Business Management, pp 171–190
43. Neeraj P, Vinay D, Deependra S (2018) Islands of interest: Mining concentrations of user search intent over e-commerce product categories. In: 2018 IEEE International Conference on Big Data, pp 3717–3722
44. Parikshit S, Mohit S, Pranam K, Cheng-Xiang Z (2018) A taxonomy of queries for e-commerce search. In: The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval, pp 1245–1248
45. Park S, Lee D. An empirical study on consumer online shopping channel choice behavior in omni-channel environment. *Telematics Inf.* 2017;34:1398–407.
46. Pazienza M, Pennacchiotti M, Zanzotto F (2005) Terminology extraction: an analysis of linguistic and statistical approaches. In: Knowledge mining, Springer, pp 255–279
47. Pierrakos D, Paliouras G. Personalizing web directories with the aid of web usage data. *Trans Knowl Data Eng.* 2010;22:1331–44.
48. Prateeti M, Anupama R, Gargi D (2018) Futse: a fuzzy taxonomy service to facilitate product search. In: Proceedings of the ACM India Joint International Conference on Data Science and Management of Data, pp 220–229
49. Raunich S, Rahm E (2012) Towards a benchmark for ontology merging. In: On the move to meaningful internet systems: OTM 2012 Workshops, Springer, pp 124–133
50. Sacco G, Nigrelli G, Bosio A, Chiarle M, Luino F. Dynamic taxonomies applied to a web-based relational database for geo-hydrological risk mitigation. *Comput Geosci.* 2012;39:182–7.
51. Sands S, Ferraro C, Campbell C, Pallant J. Segmenting multichannel consumers across search, purchase and after-sales. *J Retail Consum Serv.* 2016;33:62–71.
52. Schneider P, Zielke S. Searching offline and buying online - an analysis of showrooming forms and segments. *J Retail Consum Serv.* 2020;52: 101919.
53. Schulten E, Akkermans H, Botquin G, Doerr M, Guarino N, Lopes N, Sadeh N. The e-commerce product classification challenge. *Intell Syst.* 2001;16:86–93.
54. Sujatha R, Bandaru R, Rao R. Taxonomy construction techniques-issues and challenges. *Indian J Comput Sci Eng.* 2011;2:661–71.
55. Tarus J, Niu Z, Yousif A. A hybrid knowledge-based recommender system for e-learning based on ontology and sequential pattern mining. *Fut Gen Comput Syst.* 2017;72:37–48.
56. Tvarozek M, Bielikova M (2007) Personalized faceted navigation for multimedia collections. In: Proceedings of the 2nd International Workshop on Semantic Media Adaptation and Personalization, IEEE, pp 104–109
57. Van Rees R. Clarity in the usage of the terms ontology, taxonomy and classification. *Cib Rep.* 2003;284(432):1–8.
58. Vandic D, van Dam J, Frasincar F. Faceted product search powered by the semantic web. *Decis Support Syst.* 2012;53:425–37.
59. Verhoef P, Kannan P, Inman J. From multi-channel retailing to omni-channel retailing: introduction to the special issue on multi-channel retailing. *J Retail.* 2015;91:174–81.
60. Wang Y, Lin H, Tai W, Fan Y. Understanding multi-channel research shoppers: an analysis of internet and physical channels. *IseB.* 2015;14:389–413.
61. White A, Newman D, Logan D, Radcliffe J. Mastering master data management. Stamford: Gartner Group; 2006.
62. Xiaohui T, Yuefeng L, Ning Z. A personalized ontology model for web information gathering. *Trans Knowl Data Eng.* 2011;23:496–511.
63. Xu Z, Du D, Xu D. Improved approximation algorithms for the max-bisection and the disjoint 2-catalog segmentation problems. *J Comb Optim.* 2014;27:315–27.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.