Focused Categorization Power of Ontologies: General Framework and Study on Simple Existential Concept Expressions

Vojtěch Svátek [a,*], Ondřej Zamazal [a], Viet Bach Nguyen [a], Jiří Ivánek [a] Ján Kľuka [b] and
Miroslav Vacura [a,c],

[a] *Department of Information and Knowledge Engineering, Prague University of Economics and Business, Nám. W. Churchilla 4, 130 67 Praha 3, Czech Republic*
E-mail: {svatek, ondrej.zamazal, viet.nguyen, ivanek }@vse.cz
[b] *Department of Applied Informatics, Comenius University in Bratislava, Mlynská dolina, 842 48 Bratislava, Slovakia*
E-mail: kluka@fmph.uniba.sk
[c] *Department of Philosophy, Prague University of Economics and Business, Nám. W. Churchilla 4, 130 67 Praha 3, Czech Republic*
E-mail: vacuram@vse.cz

Abstract. When reusing existing ontologies for publishing a dataset in RDF (or developing a new ontology), preference may be given to those providing extensive subcategorization for important classes (denoted as focus classes). The subcategories may consist not only of named classes but also of compound class expressions. We define the notion of focused categorization power of a given ontology, with respect to a focus class and a concept expression language, as the (estimated) weighted count of the categories that can be built from the ontology's signature, conform to the language, and are subsumed by the focus class. For the sake of tractable initial experiments we then formulate a restricted concept expression language based on existential restrictions, and heuristically map it to syntactic patterns over ontology axioms (so-called FCE patterns). The characteristics of the chosen concept expression language and associated FCE patterns are investigated using three different empirical sources derived from ontology collections: first, the concept expression pattern frequency in class definitions; second, the occurrence of FCE patterns in the Tbox of ontologies; and last, for class expressions generated from the Tbox of ontologies (through the FCE patterns); their 'meaningfulness' was assessed by different groups of users, yielding a 'quality ordering' of the concept expression patterns. The complementary analyses are then compared and summarized. To allow for further experimentation, a web-based prototype was also implemented, which covers the whole process of ontology reuse from keyword-based ontology search through the FCP computation to the selection of ontologies and their enrichment with new concepts built from compound expressions.

Keywords: ontology reuse, class expression, focus class, categorization, OWL, ontology pattern

1. Introduction

The main motivation of providing machine-readable semantics to data on the web in the form of ontologies is that of achieving interoperability of independently built data sources and applications. For example, if the

same kind of product offered by different e-shops is semantically described using the same web ontology, comparison and automatic recommendation of these offers can be provided to customers.

Obviously, interoperability depends not only on the existence of ontologies but also on their *reuse*. Rather than coining new entities in isolation, a dataset publisher should invest into finding relevant ontologies

---

*Corresponding author. E-mail: svatek@vse.cz

and integrating their entities into the schema of the given dataset. Similarly, the designers of a new ontology should consider reusing parts of existing ontologies from the same domain. Besides the interoperability benefit, reusing a categorization structure already existing in another ontology may also save a part of the design effort.

Since the majority of ontologies is nowadays published in the same standard language, OWL [2], the reuse is easy from the technological point of view, whether the method is the direct reuse of existing ontology entities or their subsumption/equivalence mapping from the dataset schema or from the new standalone ontology. However, despite the general agreement on the benefits of ontology reuse, this best practice is not massively adhered to yet [5]. One reason might be that selecting an ontology, or a fragment of it, suitable for being reused, from a larger pool of ontologies (typically pre-selected via keyword-based search in ontology repositories) is a non-trivial task for which only recently formal methods have emerged. They mostly rely upon

- ontology/entity *popularity metrics*, such as how many instances in how many different datasets already refer to them, and
- their a priori *credibility* [21], especially, whether the ontologies are listed in catalogs such as LOV – Linked Open Vocabularies[1] [27].

However, these approaches face the 'cold start' problem. Due to rapid growth of the ontological 'ecosystem' on the web, many emerging ontologies relevant for a certain dataset (or, generally, a reuse case) might not yet have achieved significant popularity ratings. Furthermore, the fact that an ontology as whole is thematically related to the reuse case does not mean that it structurally fits well the data that is to be semantically described. All this calls for complementing such 'extrinsic' sources of evidence with 'intrinsic ones', reflecting what is in the ontology itself, beyond mere unstructured keyword matching; the ontology's axiom structures should be examined.

1.1. Working assumptions

Our proposed approach to enhancing web ontology reuse is based on four assumptions:

1. **Assumption #1** A large part of the use cases of ontologies on the web consists in *assigning* data objects to certain *categories* (with some consequences following from this assignment). Furthermore, prior to the assignment, the objects are already known to be instances of some (more general) class, to which we will refer as the *focus class* (FC).

2. **Assumption #2** The categories being assigned may not necessarily have the form of an *atomic concept* (named class) explicitly present in the ontology, but can also be *compound class expressions* (also called 'anonymous classes') that are built from atomic entities of the ontology (classes, properties and sometimes individuals) using modeling primitives of the description logic (DL) formalism underlying OWL. Informally written examples of such compound class expressions are, e.g., 'person having some child' or 'teacher or student'.

3. **Assumption #3** The number and 'quality' of the categories an ontology provides for a focus class could be summarized by a certain *aggregating metric*, which could then serve as indicator of the reusability of the ontology for a dataset containing objects semantically corresponding to this class as its instances.

4. **Assumption #4** The 'quality' of the *compound* categories is strongly correlated with the degree to which users would consider *creating a named class* (equivalent to this category) as meaningful.

The first two assumptions are supported, among other, by the findings of a study on competency questions by Ren et al., from 2014 [19]. They collected 168 competency questions (CQ) from two ontology projects, and clustered them into twelve archetypes. Of them, at least three correspond to tasks that can be characterized as sub-categorization of instances of a relatively generic class (here, pizza or software), namely:[2]

- #1: "Which [$CE1$] [OPE] [$CE2$]?" (e.g., "Which pizzas contain pork?")

---

[2]CE, CE and $CE2$ denote class expressions (in fact, named classes), OPE denotes an object property expression, and I denotes an individual (object). The numbering of the archetypes is preserved from the original publication [19]. For completeness: other archetypes than those three have to do with quantitative features of objects or with relationships between specific objects; interpreting them as sub-categorization would be thus less natural.

- #3: "What type of [*CE*] is [*I*]?" (e.g., "What type of software (API, Desk-top application etc.) is it?")
- #4: "Is the [*CE*1] [*CE*2]?" (e.g., "Is the software open source development?")

In archetype #1, *CE*1, exemplified by 'pizza', would be the focus class; the particular ontology should then allow to sub-categorize, in bulk, some pizzas as 'containing pork' (which would be represented as a compound class expression subscribing to our Assumption #2). Archetype #2 is quite similar: *CE*, exemplified by 'software', would be the focus class; the ontology should allow to sub-categorize (an individual piece of) software to some more specific class (presumably, a named one). Finally, in archetype #3 the focus class is *CE*1; the ontology should allow to verify the sub-categorization of (again, an individual piece of) software to a fixed more specific class (*CE*2).

Note that while Assumption #2 gives a value to compound class expressions, we can realistically expect that for certain reuse cases (e.g., when the categories have to be imported into a static hierarchy such as a thesaurus or product catalog on a web page) the compound class expressions have to be structurally and lexically *transformed* to named classes.

The validity of Assumption #3 obviously depends on the meaning of "(ontology) provides (categories)". Namely, an ontology is, structurally, a collection of *axioms* rather than of categories (class expressions). The generation of categories from ontology axioms, allowing for their aggregated 'counting' by a metric, will essentially be a heuristic process.

Assumption #4 brings a bottleneck: the respective notion of quality escapes an automatic empirical evaluation, as it depends on the subjective perception of users. We however believe that the direct human perception factor is indispensable for assuring the versatility of the reused schema with respect to different use cases. We could of course imagine some alternative, indirect quality measures, in particular, the distinction whether an expression would fit the *expressiveness of the query engine* to be used for querying the dataset described with the ontology. There could be however many mutually overlapping queries, and the expressions corresponding to each of them would only mildly contribute to the semantic underpinning of the dataset. Imagine for example two ontologies: ontology A has two relevant atomic concepts (subclasses of the focus class), and ontology B has four of them. Without caring for human perception, we might be tempted to con-

struct the compound categories using the conjunction operator, since conjunctive queries are the cornerstone of ontology-based data access. Ontology B would then probably achieve a far higher value of the aggregating metric, having five times more (15 vs. 3) possible non-empty conjunctions) than ontology A. It could however easily happen that none of the conjunctions would have any special meaning beyond that trivially following from its constituents, which would make this strong degree of preference for ontology B rather inadequate. For these reasons, and also due to the fact that the computational use of the reused ontologies may be rather diverse and unpredictable a priori, we view the metric computation as rather agnostic of this usage.

1.2. Motivating example

Let us now present a concrete motivation example, which we will use throughout the paper to illustrate various components of our approach.

Example 1 A used vehicle retailer website is to be enhanced with RDF descriptions of the offered vehicles. The descriptions should refer to suitable ontologi/es whenever possible, in view of achieving interoperability with (presently unknown) applications of partners, search engines and aggregators. In other words, the ontologies will be reused as parts of the schema of the dataset (or, as recently called, knowledge graph) consisting of all the structured descriptions. Now, how can we assess the potential of different existing ontologies for being reused in this case? And how will the reuse itself then take place?

- Various customers are interested in different *categories* of vehicles (in terms of technical parameters, make, operation history, etc.); the reused ontologies thus should allow to represent as many such user-demanded categories as possible. The categories can then be assigned to data items representing the corresponding vehicles within the website, this way made easier to search and browse. Note that the number of vehicle categories may not be correlated with the number of classes in the ontology. For the task of recommending a particular vehicle, categories refining the 'vehicle' concept are more relevant than others. For example, if a considered ontology *O* were a broader one covering transport in general, its capability to express categories of, e.g., traffic signs, would not be an argument in favor of its reuse in the given case.

– Ontologi/es allowing to express many 'meaningful' categories of vehicles – whether as *named classes* or as newly constructed *compound concepts* – are good candidates for becoming part of the dataset schema.

– The whole or part of the reused ontologi/es (after a merging step, if more ontologies are reused for the 'vehicle' concept) would possibly give rise to a product taxonomy published as a navigation structure of the shop website.

To name some hypothetical applications (in the context of this particular domain) where the category-centric processing of data would make sense:

– within the retailer website, the user could *navigate* over a (multi-)tree consisting of both original and constructed categories;

– the designed categories would be considered as dimensional values in an internal *sales reporting* dashboard;

– meaningful sales segments to be coherently *marketed* would be identified.

1.3. Summary of the approach

We will now leverage on the presented example in order to explain the idea of our approach in intuitive terms, prior to formalizing it in Section 2. We will present the basic principles of focus categorization power computation, explain its broader ontology reuse context, and clarify the used terminology.

Focus categorization power computation   OWL [2], being a primarily *concept-based* formal ontology language, allows to create compound class expressions of many kinds using a large collection of constructors. We define the notion of *concept expression language*[3] as a restricted formal sublanguage of OWL that only allows the creation of (certain kinds of) class expressions. The language consists of *concept expression patterns* (CE patterns), defining the structure of the expression, and of further *restrictions* (that cannot be conveniently expressed through the patterns themselves).

Depending on the chosen language, we can *build*, from the signature (i.e., set of classes, properties and individuals) of an ontology, composed class expres-

sions, for example, $\exists$ *hasParkingCamera.*$\top$ (if the language allows the existential restriction, say, through the CE pattern $\exists P.C$; $\top$ is the top concept, i.e., 'anything'), for cars with a parking camera, or, *Van* $\sqcup$ *Limousine* (if the language allows the union operator, through the CE pattern $C \sqcup D$), for cars that are either a van or a limousine. Expressions that can conform to a language $\mathcal{L}$ will be denoted as $\mathcal{L}$-expressions.

These expressions can further subcategorize entities of which we already know their *focus class*, i.e. *Car* (or, *Vehicle* or similar). The conjunction of the focus class *FC* and the subcategorizing expression will be denoted as an $\mathcal{L}$-category of *FC*. In our case, one of the $\mathcal{L}$-categories of *Car* could be *Car* $\sqcap \exists$ *hasParkingCamera.*$\top$.

A concept expression language and the signature of the ontology alone are not sufficient as input for building 'meaningful' categories. In most cases, the vast majority of the combinations of entities (fitting the language structure) would be nonsensical, as the entities would be mutually irrelevant. Therefore, we need one more ingredient: heuristic *syntactic patterns* relying on templated axioms. Upon successfully matching such a pattern (called FCE pattern) to the ontology as collection of axioms, a class expression will be constructed, with the help of a mapping function specific for the given FCE pattern. An example of an FCE pattern, later formally defined in Section 4, Equation (7)), is that expecting the existence of a domain axiom for a certain property, stating that the subject of this property is the focus class. In our example, an axiom might state that the domain of the *hasParkingCamera* property is *Car*. Consequently, under some conditions, the CE $\exists$ *hasParkingCamera.*$\top$ would be generated through the mapping function that is part of the FCE pattern.

The ability of the ontology to subcategorize the focus class, the *focus categorization power* (FCP), would ideally be expressed as a sum of *weights* (degrees of 'meaningfulness') of all individual class expressions that have been generated. For example, as humans we can judge that *Car* $\sqcap \exists$ *hasParkingCamera.*$\top$ would be a much more meaningful category of *Car* than *Car* $\sqcap \exists$ *hasEngine.*$\top$ (since a car is nearly *always* sold together with its engine, it is useless to explicitly distinguish it) or *Car* $\sqcap \exists$ *hasSeller.Child* (since a car is probably *never* sold by a child). Such a judgement is however beyond the reach of a machine – at least, unless equipped with cutting-edge common-sense reasoning, which would be a sledgehammer for a nut when merely trying to suggest ontologies for reuse.

---

[3] Formally we should use the term 'OWL concept expression language'; we omit the 'OWL' attribute for brevity, since we do not go beyond the limits of OWL in any part of this research.

Therefore we propose to approximate the 'meaningfulness' – or, in more practical terms, *reusability* – of *individual* class expressions with the 'average' reusability of their underlying *concept expression patterns* (bundled with their corresponding syntactic FCE patterns). In this respect, $\exists hasParkingCamera.\top$ and $\exists hasEngine.\top$, having the same underlying structure (concept expression pattern), will be counted with the same weight, but $\exists hasSeller.Child$ will be counted with a different weight in the FCP calculation, since it has a specific class (*Child*) as the restriction filler, instead of the top concept ($\top$).

Ontology reuse process   Considering the used-cars example as an end-to-end scenario, the (re)usability analysis of a set of ontologies should first seek in them, using lexical methods, classes expressing the general notion of 'vehicle'. In each ontology, it will make such a class the *focus class*, and perform the FCP calculation with respect to it. A high value of the FCP will indicate the suitability of the given ontology for becoming a part of the schema of the website's data. Then the *actual reuse* will take place, in which the previously generated compound class expressions could become the definitions of new named classes. In the navigation structure of the used-cars website, the compound concepts thus would not have to be served as logical formulas, but could be *transformed* (ideally, in a partially automated way) into regular noun phrases that would, explicitly or implicitly, correspond to *named OWL classes* – e.g., *Cars that have parking camera*.

Terminological note   In the context of introducing the terminology to be employed in the rest of the paper, we should also make explicit our usage of (closely related) terms *concept*, *class* and *category*. For the first two, we subscribe to the common practice within the semantic web community, where both are used nearly interchangeably, with 'concept' being more prominent within the description logic literature, while 'class' being the official term used in OWL specifications. 'Concept', in more general terms, can also be the calling for the *intension* of the class, i.e., for its underlying meaning independent of its extension (set of instances). To maintain the link with OWL specifications, we denote the individual expressions as 'class expressions'; however, in the newly introduced terms ('concept expression pattern', 'concept expression language') we use the 'concept' alternative, to hint at the important role of human perception in our research. Finally, as regards the term 'category', we endow it with particular flavor: as we see, a category is always related to a more general class/concept (i.e., the focus class) which it *subcategorizes*. Additionally, we associate a category with a certain *concept expression language* $\mathcal{L}$ in which it can be expressed (this is why we will formally introduce it as an $\mathcal{L}$-category in Section 2).

1.4. Research contributions and paper structure

The research presented in this paper aims to bring the following contributions:

1. Introduce the notion of *focused categorization power* of ontologies as a fundamentally new concept in analyzing ontologies (particularly relevant in ontology reuse scenarios, but possibly even beyond), position it with respect to *related research*, and provide it with a *formal underpinning* and an *algorithmic solution*.
2. Demonstrate this ontology analysis approach using a particular *concept expression language* (and an associated collection of FCE patterns), which is simple but sufficiently rich to show various aspects and challenges of the approach.
3. The demonstration is not merely descriptive, but also features *empirical* and *cognitive* studies, which on the one hand produce tentative *pattern weighting* needed for operational FCP computation, and on the other hand provide some general insights potentially interesting for the ontology engineering community in general.
4. Finally, further experimentation is made easier for other researchers by making available a prototype *implementation* of the whole approach, including a convenient GUI.

The present paper is an evolution of a previous conference paper [23], which provided a brief explanation of the notion of FCP, informally proposed a concept expression language with FCE patterns, and provided the results of a first cognitive experiment. The present paper however extends the previous one along numerous axes, and the majority of its volume consists of entirely new content:

– The *core FCP model* has been completely reworked and extended by a number of notions, thus gaining more formal rigor.
– The survey on FCE pattern *occurrence in ontology collections*, only present in an online addendum to the previous paper [23], is now an integral part of the paper, and has been re-run to provide fresher statistics. Its results are now also more thoroughly discussed.

– A study on the presence of *compound class expressions* inside axioms has been newly added.
– The *cognitive experiment* has been repeated on a different data sample, with more guidance provided to the experimental subjects, and with additional meta-data collection.
– The different analyses / experiments in the paper are now framed by a *comparative analysis*.
– The prototype *implementation* (named OReCaP), now available both as a running instance and in source code, is an entirely new result.[4]

The rest of the paper is structured as follows. Section 2 provides a formalization of the focused categorization power framework (outlined in intuitive terms in Section 1.3). Section 3 introduces a CEL having suitable properties for an initial study. Section 4 complements this language with FCE patterns. Section 5 surveys the occurrence of concept expression patterns in ontology axioms. Section 6, analogously, surveys the occurrence of FCE patterns in the Tbox of ontologies. Section 7 describes a series of cognitive experiments in which humans provided an assessment of 'meaningfulness' of class expressions belonging to different patterns. Section 8 provides a comparison of those complementary surveys and experiments and a summarizing discussion. Section 9 describes a tentative operationalization of the results from the analyses, and explains the functionality of the implemented OReCaP prototype. Section 10 reviews some related methods and projects. Finally, Section 11 wraps up the paper and outlines the directions for future work.

2. General framework of focused categorization power

The aim of this section is to formally underpin the whole approach as well as to motivate the empirical analyses to which most of the remainder of the paper is devoted. We will proceed from the notion of *concept expression language* to the notion of $\mathcal{L}$-*category*, as a concept (class) expression refining a (focus) class and conforming to some concept expression language, and then to the central notion of *focused categorization power* (FCP). Finally, we analyze what sources may inform the computation of the FCP through a *weighting* mechanism and how the FCP can be approximated by *matching FCE patterns* in the ontology.

---

[4]Its early version was presented as a short demo paper [16]; the current version however has substantially enhanced functionality.

2.1. Concept expression language

The use of the syntactic constructors in OWL can be restricted in different ways, producing a formal system in logic often called a fragment or sublanguage. There is a number of decidable fragments of description logics [10]. The so-called OWL Profiles as sublanguages of OWL defined in the current OWL 2 standard are examples of such fragments [1]; however, non-standardized restricted sublanguages of OWL may be also be useful for particular tasks, as here.

The notion of concept expression language, in our terms, is based upon a set of concept expression *patterns* constructed over a set of variables together with a set of allowed *substitutions* for those variables. Given a particular ontology signature, a concept expression language generates a set of concept expressions by applying all allowed substitutions to all patterns.

Let us first recall or introduce a few preliminary notions. A *signature* is a triple $\mathcal{S} = (\mathcal{C}_\mathcal{S}, \mathcal{P}_\mathcal{S}, \mathcal{I}_\mathcal{S})$ of sets of classes $\mathcal{C}_\mathcal{S}$, properties $\mathcal{P}_\mathcal{S}$, and individuals $\mathcal{I}_\mathcal{S}$, collectively called *entities*. The *signature of an ontology O* is the signature $sig(O) = (\mathcal{C}_O, \mathcal{P}_O, \mathcal{I}_O)$ where $\mathcal{C}_O, \mathcal{P}_O, \mathcal{I}_O$ are respectively the sets of all classes, properties and individuals occurring in O (cf. [2, Sect. 5]). Note that the punning feature of OWL 2 allows these sets to be non-disjoint. By $\mathcal{PE}_\mathcal{S}$ and $\mathcal{CE}_\mathcal{S}$ we denote, respectively, the sets of property expressions and class expressions (CEs) over the signature $\mathcal{S}$, defined in the standard way [2, Sects. 6 and 8].

In order to bring in the notions of variable and substitutions, let us introduce a special signature $\mathcal{V} = (\mathcal{C}_\mathcal{V}, \mathcal{P}_\mathcal{V}, \mathcal{I}_\mathcal{V})$ such that the sets $\mathcal{C}_\mathcal{V} = \{C, C_1, C_2, \ldots\}$, $\mathcal{P}_\mathcal{V} = \{P, P_1, P_2, \ldots\}$, $\mathcal{I}_\mathcal{V} = \{i, i_1, i_2, \ldots\}$ are countable, infinite, and mutually disjoint. The entities in $\mathcal{C}_\mathcal{V}$, $\mathcal{P}_\mathcal{V}$, $\mathcal{I}_\mathcal{V}$ will be, respectively, called *class variables*, *property variables*, and *individual variables*, and collectively we will call them *variables*. Class expressions $\mathcal{CE}_\mathcal{V}$ and property expressions $\mathcal{PE}_\mathcal{V}$ over $\mathcal{V}$ are then called *concept expression patterns* (CE patterns, for brevity) and *property expression patterns*, respectively.

Further, given a signature $\mathcal{S}$, any function σ that maps class variables to class expressions over $\mathcal{S}$ (i.e., $\sigma(C) \in \mathcal{CE}_\mathcal{S}$ for each $C \in \mathcal{C}_\mathcal{V}$), property variables to property expressions over $\mathcal{S}$ ($\sigma(P) \in \mathcal{PE}_\mathcal{S}$ for each $P \in \mathcal{P}_\mathcal{V}$), and individual variables to individuals in $\mathcal{S}$ ($\sigma(i) \in \mathcal{I}_\mathcal{S}$ for each $i \in \mathcal{I}_\mathcal{V}$) is called a *substitution*. Let $\Sigma_\mathcal{S}$ denote the set of all substitutions for a signature $\mathcal{S}$. Substitutions can easily be inductively extended to property and concept expression patterns in the obvious way, e.g., $\sigma(p^-) = \sigma(p)^-$,

$\sigma(c_1 \sqcup c_2) = \sigma(c_1) \sqcup \sigma(c_2)$, $\sigma(\exists p.c) = \exists \sigma(p).\sigma(c)$, $\sigma(\{i_1, \ldots, i_n\}) = \{\sigma(i_1), \ldots, \sigma(i_n)\}$, etc. (cf., Baader & Narendran [4]).

With preliminaries in place, we may proceed to our definitions.

Definition 1 *A concept expression language (CEL) is a pair $\mathcal{L} = (\mathcal{CP}_\mathcal{L}, \mathcal{R}_\mathcal{L})$, where*

- *$\mathcal{CP}_\mathcal{L}$ is a set of concept expression patterns (CE patterns), and*
- *$\mathcal{R}_\mathcal{L}$ is a substitution-restricting function mapping every signature $\mathcal{S}$ to a subset of $\Sigma_\mathcal{S}$, the set of allowed substitutions.*

One example of a CEL may be the trivial language of named concepts $\mathcal{L}_{nam} = (\{C\}, \mathcal{R}_{nc})$ where the subtitution-restricting function $\mathcal{R}_{nc}$ only allows a named (atomic) class to be substituted for C. Thus for every signature $\mathcal{S}$, the set $\mathcal{R}_{nc}(\mathcal{S})$ contains substitutions that map the class variable C to a member of $\mathcal{C}_\mathcal{S}$.

Another CEL example is the language of existential restrictions to the top concept $\mathcal{L}_{extop} = (\{\exists P.\top\}, \mathcal{R}_{np})$ where for each signature $\mathcal{S}$ the set $\mathcal{R}_{np}(\mathcal{S})$ contains only substitutions that map the property variable P to a named property of $\mathcal{S}$.

For clarity, we explicitly define the notion of class expressions generated by a CEL:

Definition 2 *Given a CEL $\mathcal{L} = (\mathcal{CP}_\mathcal{L}, \mathcal{R}_\mathcal{L})$ and a signature $\mathcal{S}$, any class expression $\sigma(c)$ obtained by applying a substitution $\sigma \in \mathcal{R}_\mathcal{L}(\mathcal{S})$ to a CE pattern $c \in \mathcal{CP}_\mathcal{L}$ is an $\mathcal{L}$-expression in $\mathcal{S}$. We say that $\sigma(c)$ conforms to (also matches) the CE pattern c.*

2.2. $\mathcal{L}$-categories and focused categorization power

Only some of the $\mathcal{L}$-expressions in an ontology signature are relevant for us: those sub-categorizing the given *focus class*; we denote them as *$\mathcal{L}$-categories* for this class. This will allow us to formulate the notion of FCP, at the abstract (non-operational) level first.

Definition 3 *Let $\mathcal{L} = (\mathcal{CP}_\mathcal{L}, \mathcal{R}_\mathcal{L})$ be a CEL. Let us assume an OWL ontology O with signature $sig(O) = (\mathcal{C}_O, \mathcal{P}_O, \mathcal{I}_O)$ and a named class $FC \in \mathcal{C}_O$. An $\mathcal{L}$-category[5] for FC in O is then any class expression $Cat = FC \sqcap D$ where D is an $\mathcal{L}$-expression in $sig(O)$.*

_____

[5] For simplicity, we will sometimes simply write 'category' rather than $\mathcal{L}$-category throughout the paper, provided the language $\mathcal{L}$ is either clear from the context or irrelevant.

The expression D is then called a differentiating class expression *(DCE) of the $\mathcal{L}$-category.*

Within this paper we will sometimes informally refer to a DCE as to the *subcategory* of the class *FC*. Note, however, that DCE is a mere syntactic notion, as for some *FC* and *D* it would not hold that $FC \sqcap D$ would 'differentiate' among the instances of *FC* – i.e., as we will discuss later, $FC \sqcap D$ may not be extensionally a *proper* subconcept (subcategory) of *FC*.

The capability of *O* to finer subcategorize the resources that are already known to belong to *FC* can be, at an abstract level, expressed in terms of the number as well as 'quality' of $\mathcal{L}$-categories for *FC* in *O*, for a suitably selected $\mathcal{L}$. We will denote the measure of this capability as *focused categorization power* (FCP) of *O* with respect to *FC* over $\mathcal{L}$; the class *FC* plays the role of a *focus class*. Formally:

$$FCP(O, FC, \mathcal{L}) = \sum_{Cat \in Cats}^{k} w(Cat, O, FC) \qquad (1)$$

where *Cats* is the set of all $\mathcal{L}$-categories for *FC* in *O* unique under the inferential closure (i.e., one representative of each set of equivalent categories is present in *Cats*) and *w* is a *weight function* assigning every *Cat* a weight from $[0, 1]$, expressing its 'quality' as a category of *FC* in *O*.

2.3. Weight function desiderata

Still remaining at the (non-operational) 'guideline' level, we should discuss some desiderata of the weight function just introduced. Intuitively, leveraging on some of the assumptions from Section 1.1, an ideal function $w(Cat, O, FC)$ should take into account the following four characteristics of an $\mathcal{L}$-category *Cat*:

Absolute size The weight *w* should be 0 if the extension of *Cat* is an empty set (*Cat* is an unsatisfiable concept). *w* should also be low if the extension of *Cat* is a singleton (or, very small) set.

Relative size The weight *w* should be 0 if *Cat* is equivalent to *FC*. On the other hand, *w* should be highest if *Cat* contains a 'reasonable' proportion of the instances of FC, neither extremely high nor extremely low. For example, a category with around 10 instances may be quite meaningful for a focus class with a few dozen of instances (say, the DCE `EuropeanMonarchy` wrt. the focus class `EuropeanCountry`), but less so for a focus class

with millions of instances (say, a DCE expressing the notion of 'village next to a spacecraft launch location' wrt. the focus class `Village`).

Expression complexity The weight w should increase with the decreasing complexity of *Cat* as DL expression. The rationale is that a simpler category can be better decoded by people and more conveniently displayed. The expression complexity is influenced by the choice of the CEL. In this paper we will considering a rather restricted CEL, in which the expression complexity will not be of concern for any of the CE patterns; therefore we will not explicitly consider it in the operational FCP computation.

Semantic coherence The weight w should be lowered if a category is assembled from thematically unrelated entities. An example derived from the DBpedia ontology (we also refer to it in Section 7.1) is 'person beatified in a wine region'; it is unlikely that such a concept spanning across the religious and agricultural domain would be of high interest in any of them. However, this characteristic is least eligible for automatic assessment of all four,[6] we thus do not elaborate on it here either.

Checking whether the *size* of a category *Cat* (whether absolute, or relative vs. *FC*) is such that leads to $w(Cat, O, FC) = 0$ is to some degree possible via a DL reasoner. This characteristic is thus the only one we will explicitly consider in the FCP computation. Let us provide a preliminary illustration of such a consideration, on the previously introduced language of *named classes*, $\mathcal{L}_{nam}$. Each $\mathcal{L}_{nam}$-category is a conjunction of the given focus class *FC* and another named class *D*. Intuitively, $w(FC \sqcap D, O, FC)$ should be 0 for such *D* that either $FC \sqcap D = \emptyset$ (i.e., for classes disjoint with *FC*) or $FC \sqsubseteq D$ (i.e., for superclasses of *FC*) due to violating the hard constraint on absolute/relative size. On the other hand, w should be high (presumably even equal to 1) for the asserted (even transitive) subclasses of *FC*, since these are the concepts viewed as proper subcategories of *FC* by the ontology designer/s. Note however that w might deserve to be non-zero even for some other non-disjoint classes. For example, `Student` $\sqcap$ `Teacher` can be a valid category of the FC `Student` in an academic ontology, since some graduate students may also have a teacher contract; obviously, none of the classes would be a subclass of the other.

---

[6]Presumably, graph-based metrics relying on the number of different paths connecting the constituent entities might be applied.

2.4. Focused category extraction pattern

Before proceeding to an operational elaboration of the category weight function in FCP computation, we need to return to the problem of category generation. While the CE patterns defining a given CEL allow for substitutions producing class expressions to be 'counted' in the FCP computation, such expressions cannot be directly extracted from the OWL code, which consists of (TBox, and sometimes ABox) *axioms* and not of standalone expressions. We will thus introduce the general notion of *focused category extraction pattern* – a pattern for transforming OWL axioms occurring in ontologies to class expressions to be used as the DCE of categories of a focus class.

The definition of a focused category extraction pattern relies on OWL *axiom templates* (axioms containing variables) that can be matched to an OWL ontology through a unification substitution (returning a set of variable bindings with ontology entities). In order to eliminate undesirableg matches without burdening the axiom templates with complex (negated) descriptions of exceptions, *validity constraints* are furthermore introduced. Finally, the generation of CEs is carried out by *mapping functions*.

Definition 4 *A* focused category extraction pattern *(an FCE pattern, for short) is a 5-tuple* $p = (c, Vars, Tpl, Cons, mf)$ *where*

- – *c is a* CE pattern.
- – *Vars is a sequence of* mapping variables.
- – *Tpl is a set of conjunctively*[7] *interpreted axiom templates, namely, OWL axioms (possibly) containing variables in the position of entities. One distinguished variable is (usually, though not mandatorily) the* FC-variable, *which is to be bound to the focus class FC when matching the pattern. Every variable from Vars has to appear in Tpl at least once, and there can also be other (non-mapping) variables in Tpl not present in Vars.*
- – *Cons is a set of FCE pattern validity* constraints.
- – *mf is a* mapping function, *which has as input a tuple of entities (from the matched ontology) to which the variables from Vars are bound, and returns a CE conforming to c, provided all the constraints from Cons are satisfied by the tuple.*

---

[7]For simplicity we only consider conjunctions of axioms. Disjunctions could be expressed by means of multiple FCE patterns for the same CE pattern, if needed.

Considering we have two kinds of patterns (CE and FC), which both abstract over some OWL structures, a question naturally arises why these two could not be more integrated or based one upon the other. There are however two factors that would make such an integration cumbersome or impossible:

- First, while the variables in CE patterns are strictly *typed* (in the sense that the abstract signature $\mathcal{V}$ consists of sets of variables for classes, properties and individuals that are pairwise disjoint), the FCE patterns should be versatile wrt. possible use of *punning* in the ontology: across multiple axioms, the same identifier might, e.g., once refer to an individual and once to a class.
- Second, the role of the FCE patterns is not merely in matching the OWL structures but also in constructing class expressions based on the match. The CE pattern of the constructed expressions will not always (probably, just rarely) correspond to the structure of expressions explicitly present in the axioms matched by the FCE pattern.

For these reasons, we keep the two notions as completely separate. In the rest of the paper we will carefully distinguish which type of pattern is being discussed (using the 'CE' and 'FC' acronyms) unless the type of pattern is entirely clear from the context.

2.5. Hypothetical sources of category weight

Assignment of a weight w to a certain $\mathcal{L}$-category $Cat = FC \sqcap D$ while computing $FCP(O, FC, \mathcal{L})$ in (1) could be hypothetically based on a number of different sources, possibly in combination; note that the sources refer to both CE patterns and FCE patterns, which we have previously distinguished:

1. *FCE pattern axiom template* matching, applied on the context of *FC* and (constituent entities of) *D* in *O*. If the template *Tpl* is matched, a CE can be generated via the mapping function *mf* (and possibly obtain a non-zero weight).
2. *Inferential relationships* between *FC*, the entities from which *D* consists, and possibly other entities in *O*. The validity of such inferential relationships can be checked via constraints declared in the *Cons* element of an FCE pattern; if the constraints do not hold then *w* should be shrunk.

   - As pointed above for the specific case of *D* being a named class, but valid generally, if $FC \sqcap D = \emptyset$ or $FC \sqsubseteq D$ then *w* should be 0.

   - Moreover, if $|FC \sqcap D| = 1$ then $FC \sqcap D$ is not a very useful category either. This happens, e.g., with *D* being $\exists P.\{i\}$ such that *i* is a particular individual and *P* is an inverse-functional property (inferentially yielding $|D| \leq 1$).

3. Frequency of instantiation of *Cat* in some *RDF dataset/s* (Abox). The weight could be arithmetically derived from the frequency.
4. Likelihood that the *CE pattern* $c \in \mathcal{CP}_{\mathcal{L}}$ to which *D* conforms produces a meaningful category on average. This likelihood could be inductively aggregated from the weights of a representative collection of CEs corresponding to CE pattern *c*, and then deductively applied to *D*.

Let us demonstrate these sources on the vehicle domain of our motivating example.

Example 2 A considered ontology contains in its signature a focus class *Car*[8] and a property *hadAccident*. In a CEL containing the existential restriction $\exists$ and the top concept $\top$ (such as $\mathcal{L}_{extop}$ from in Section 2.1) the category *Car* $\sqcap \exists$*hadAccident*.$\top$ can be constructed. The DCE of this category conforms to the CE pattern $\exists P.\top$. Through the lenses of the four sources:

1. *FCE pattern axiom template*: Let the property *hadAccident* have *Car* as its domain. An *FCE pattern* based on the *rdfs:domain* property (see Section 4 for details) is then matched in the ontology TBox. The rationale of heuristically linking this FCE pattern to the CE pattern $\exists P.\top$ can be illustrated as follows: since everything having had an accident is inferred to be a car thanks to the domain axiom, we are at least (almost) sure that there exist cars that have had an accident; there is also a chance (though not a guarantee) that there might also exist cars that have not have had one yet. The pattern thus suggests the category as possibly meaningful.
2. *Inferential relationships*: If the axioms of the ontology are consistent with reality, the *reasoner* would not infer any prohibitive statement, since the set of cars that have had an accident is neither equivalent to the set of all cars, nor is it empty or a singleton. (Hypothetically, if all cars

---

[8]For easy readability of the DL formulas, we will mostly use the simple DL notation without IRI prefixes in the examples. The namespace will be irrelevant for artificial examples and clear from the context in real-world examples.

in the universe have had an accident, $Car \sqcap \exists hadAccident.\top$ would not be a meaningful category of Car; however, a carefully modeled ontology should then contain the respective existential axiom as a necessary condition on the Car class: $Car \sqsubseteq \exists hadAccident.\top$). The source thus does not 'forbid' the suggested category.

3. *RDF datasets (Abox)*: Let us assume there is already a publicly available car *dataset* published using this ontology, which contains 2500 instances of *Car*, of which 1000 appear as subject in at least one triple with predicate *hadAccident*. The category, if translated to a SPARQL[9] query to this dataset,[10] is instantiated 1000 times, which looks like a solid absolute size, and also the relative size wrt. the focus class frequency ($1000/2500 = 0.4$) is fine. Even this third source should thus have a positive effect on the weight.

4. *CE pattern likelihood*: As we partially verify in Section 7, the CE pattern $\exists P.\top$ is a priori not as likely to produce good categories as are some other CE patterns. This would have led to a weight decrease; however, in this case, the support of the *specific* expression by Abox data should overrule the negative impact of the mere *CE pattern* this expression conforms to.

2.6. Weight of individual category vs. of CE pattern

As suggested in the previous subsection, the ultimate weight of a category can be influenced both by the factors specific for the particular category and by the CE pattern type of the DCE of this category. The category weight computation (for the sake of FCP computation, see Equation 1, i.e., still at the non-operational level) could be possibly decomposed as

$$w(FC \sqcap D, O, FC) =$$
$$w^I(FC \sqcap D, O, FC) \otimes w^P(pattern(D)) \quad (2)$$

where

– w^I is an individual weight function for the given category (presumably based on the first three sources from the list in Section 2.5: FCE pattern axiom templates, inferential relationships, and possibly Abox occurrence),

– *pattern*(D) is the CE pattern that the (differentiating) class expression matches,

– w^P is the CE pattern weight function (based on the fourth source from the list in Section 2.5: likelihood of the CE pattern producing a meaningful category), and

– $\otimes$ is a function for combining the two partial weights into a single weight of the category.[11]

As regards the CE *pattern* weight w^P, we are aware of three empirical sources, in turn, which might provide a *quality estimate* of a certain CE pattern (and thus, indirectly, the individual categories in which the DCE would match this CE pattern):

– *Humans*, who can assess concrete expressions conforming to this CE pattern as less or more meaningful subcategories of a focus class.

– RDF *datasets* instantiating expressions of this pattern (under the inferential closure) with smaller or bigger numbers of Abox statements.

– Existing ontology *axioms* in whose right-hand sides[12] (RHS) an expression matching the given CE pattern appear less or more often.

Data collected from these sources then have to be *aggregated* across the different expressions matching the same CE pattern; the human feedback also has to be aggregated over the participating human assessors. All three sources also have their biases and limitations: *human assessment* is generally subjective, *RDF data* may not yet be available for newer ontologies, and the presence of *axioms with a compound RHS* is biased by the higher/lower tendency of different communities to axiomatize ontologies; the design rationale for the axioms (intended for DL reasoning) may also differ from the usefulness criteria of categories meant to subcategorize a focus class.

---

[9]https://www.w3.org/TR/sparql11-overview/
[10]SELECT DISTINCT ?c {?c a :Car; :hadAccident []}

[11]As we exemplified (on 'cars with an accident') in Sect. 2.5, the presence of (trustworthy) Abox data should overrule the assumptions made based on the specific Tbox information (within $w^I(FC \sqcap D, FC)$) or by the CE pattern (within $w^P(pattern(D))$). A formula truly considering the *Abox* should thus grant it a more prominent position. However, since we resign on the impact of the Abox in the next elaboration of the same formula (Eq. 3), we may afford such a simplification.

[12]By the RHS we mean the syntactic RHS as indicated in the code. For the overwhelming majority of axioms inside the OWL code available in ontology documents, the syntactic LHS is a named class. Although the OWL grammar allows for general concept inclusions, having a compound concept on the left-hand side (see http://ontogenesis.knowledgeblog.org/1288/), its usage is currently marginal. For example, the 2017 snapshot of the LOV catalog only contained three ontologies featuring them.

Orthogonally, we can also aggregate the occurrence counts of *FCE patterns* in ontologies. This fourth source, in turn, will indicate the upper bound of the *quantity* of categories that can be obtained for a focus class using such patterns.

In different sections of the paper we will empirically investigate three of the sources identified in this subsection: the CEs in axioms in Section 5, the FCE patterns in Section 6, and the human assessment of FC+category pairs in Section 7 (thus only deferring the Abox analysis to later research), and discuss their potential and limitations in more detail.

To avoid any mismatch of the presented four 'weight sources' list with the (also four) 'weight sources' from Section 2.5, note that the sources from Section 2.5 are applied 'deductively', to estimate the weight of a particular category, while the sources in this section serve for 'inductive' derivation of the (average) weight pertaining to a whole CE pattern.

2.7. FCP approximation via FCE patterns

We will now specifically consider the situation when the Abox information on the usage of a specific category is either unavailable or unreliable, and we thus have to derive the category weight from the *ontology* alone. Under such conditions, the individual weight function w^I is likely to be Boolean: the FCE pattern axiom templates *Tpl* are either matched or not, and, similarly, the verification of constraints *Cons* over the ontology inferential closure will typically either judge the category as useless (unsatisfiable, or with a singleton extension, or equivalent to the FC) or not. Under the assumption of a *Boolean* w^I, we can express the overall weight of a category as a simple product:

$$w(FC \sqcap D, O, FC) =$$
$$w^I_{01}(FC \sqcap D, O, FC) \cdot w^P(pattern(D)) \quad (3)$$

which will yield

- 0 if $w^I = 0$, and
- $w^P(pattern(D))$ if $w^I = 1$.

Let us then assume that we distinguish among k CE patterns allowed in the CEL $\mathcal{L}$, $\mathcal{CP}_{\mathcal{L}} = \{c_1, \ldots, c_k\}$, each associated with a weight via the CE pattern weight function w^P. Let each CE pattern c_i be also associated with a set of FCE patterns $PS_i = \{p_{i1}, \ldots, p_{ij_i}\}$ that allow to construct the corresponding categories from the axioms of ontologies. Then we can express

the *approximate FCP* of an ontology O for a focus class FC over $\mathcal{L}$ with respect to the (set of) FCE pattern sets $PS = \{PS_1, \ldots, PS_k\}$ as the weighted sum of the FCE pattern occurrence counts in O:

$$\widehat{FCP}(O, FC, \mathcal{L}, PS) =$$
$$\sum_{i=1}^{k} Occ(PS_i, FC, O) \cdot w^P(c_i) \quad (4)$$

where $Occ(PS_i, FC, O)$ is the sum of occurrences in O for all FCE patterns $p_{ij} \in PS_i$; we speak about an occurrence of an FCE pattern $p = (c, Vars, Tpl, Cons, mf)$ when an axiom template $tpl \in Tpl$ is matched in O *and* the constraints from *Cons* hold.

Note that in terms of Eq. 3, the multiplication by $w^P(c_i)$ (the CE pattern weight) is now performed over the *counts* of FCE pattern occurrences and not over the individual category weights, but the final weighted sum will be the same as if the weights of all individual categories produced by Eq. 3 were summed according to formula Eq. 1. Although the formulas for *FCP* (Eq. 1) and $\widehat{FCP}$ (Eq. 4) appear dissimilar, their intention is the same; $\widehat{FCP} = FCP$ would even hold under (obviously, strong) assumptions that

- each CE weight were equal to its CE pattern weight, and,
- the FCE patterns were 'perfect' in the sense of generating exactly those CEs producing a non-zero weight (if used as the DCE of a category).

2.8. Pattern-based CE extraction and approximate FCP computation algorithm

The computational process leading to extraction of class expressions using FCE patterns is rather straightforward with respect to the general description of the approach in Section 2.7, see the algorithm pseudo-code in Listing 1. Each call of the function **match**(O, FC, p_{ij}) returns a class expression generated by a new match of p_{ij} in O respective to the given focus class FC, and taking into account the constraints *Cons* in the pattern; when no new CE can be generated, the function returns *false*. As a result, a set *CE* of pairs "CE pattern, CE" is returned. The only additional step is the pruning of mutually equivalent CEs, which should be only counted once (lines 14–31). First, for each CE from *CE*, an auxiliary named class is generated (by function **newIRI**) and stored, and an axiom defining this auxiliary class as equivalent to the given

CE is added to an auxiliary version of the ontology (line 19). In the deductive closure of this auxiliary ontology, equivalence axioms between pairs of auxiliary classes are then sought (line 24), and for each such pair, only one of the CEs is retained. The choice of the CE to be discarded could be based on the weight of the CE patterns:

$$\textbf{discard}(c_i, ce_i, c_j, ce_j) = ce_i \ \ if \ \ w^P(c_i) < w^P(c_j)$$
$$= ce_j \ \ otherwise$$

For completeness, we also include the approximate FCP computation algorithm, as Listing 2. It simply amounts to summation of CE pattern weights w^P over the collected class expressions.

```
1  Input: O, FC, PS = {PS₁,...,PSₖ}
2  Output: CE
3  Procedure:
4  CE := ∅
5  for each PSᵢ
6   for each pᵢⱼ ∈ PSᵢ,
7    pᵢⱼ = (c, Vars, Tpl, Cons, mf)
8    do while match(O, FC, pᵢⱼ) ≠ false
9     ce := match(O, FC, pᵢⱼ)
10    CE := CE ∪ {(c, ce)}
11    enddo
12   endfor
13  endfor
14 Aux := ∅
15 O_aux := O
16 for each (c, ce) ∈ CE
17  aux := newIRI
18  Aux := Aux ∪ {aux}
19  O_aux := O_aux ∪ {aux ≡ ce}
20 endfor
21 O_auxc := closure(O_aux)
22 for each auxᵢ ∈ Aux
23  for each auxⱼ ∈ Aux \ {auxᵢ}
24   if {auxᵢ ≡ auxⱼ} ⊑ O_auxc then
25    find auxᵢ ≡ ceᵢ, auxⱼ ≡ ceⱼ in O_aux
26    find (cᵢ, ceᵢ), (cⱼ, ceⱼ) in CE
27    ce_d := discard(cᵢ, ceᵢ, cⱼ, ceⱼ)
28    CE := CE \ {(c_d, ce_d)}
29    Aux := Aux \ {aux_d}
30   endif
31  endfor
32 endfor
```

Listing 1: Pattern-based CE extraction algorithm

```
1  Input: CE
2  Output: F̂ĈP
3  Procedure:
4  F̂ĈP := 0
5  for each (c, ce) ∈ CE
6   F̂ĈP := F̂ĈP + w^P(c)
7  endfor
```

Listing 2: Computation of approximate FCP

2.9. *Summary and links to the remainder of the paper*

With the algorithm, we now have a general operational framework allowing us to approximately calculate the FCP of an ontology with respect to a focus class, relying purely on the axioms from the ontology.

The framework still needs to be instantiated with three components:

1. a concrete CEL, providing the CE patterns,
2. sets of FCE patterns, linked to these CE patterns via mapping functions, and, finally,
3. weights of these CE patterns.

In Section 3 we provide a CEL that is simple to work with but still sufficiently rich to demonstrate various aspects of the method. In Section 4 we suggest a set FCE patterns for its CE patterns. Finally, tentative CE pattern weights are deduced from the experiments in Section 7, while the preceding sections, 5 and 6, provide some empirical insights on the proposed CE and FCE patterns, respectively.

3. Simple existential CEL: $\mathcal{L}_{SE}$

Having explained the whole proposed framework of FCP computation, let us now return to its most essential element: the concept expression language (CEL) that determines what kinds of CEs are considered for the differentiating class expressions (DCEs) specializing the focus class.

As set up by the Assumption #2 from the introduction (importance of both atomic and compound CEs for the FCP), for meaningful analysis of the FCP computation landscape we have to combine some variant of the language of named concepts ($\mathcal{L}_{nam}$ from Section 2) with a CEL involving some DL concept constructor. For the purposes of the first proof of concept we sought a 'small' CEL well tractable, both by hu-

mans and machines, in the empirical investigations. Compared to other common types of concept constructors, the *simple existential restriction*, $\exists P.C$ where only atomic classes can be substituted for C, has some nice features in this respect – though, admittedly, subjective (we are aware that similar arguments could also be found for other candidate 'starter' CELs):

- The *interplay* between properties and classes, which are prone to be viewed separately in some ontology design environments, e.g., Protégé, may intuitively give rise to concepts 'interesting to look at', whatever their ultimately judged contribution to the FCP will be.
- The number of matching CEs only grows linearly with the number of classes $|\mathcal{C}_O|$ and number of properties $|\mathcal{P}_O|$ in the ontology (at most it can reach $|\mathcal{P}_O|.|\mathcal{C}_O|$).
- The existence of a property assertion (witnessing the validity of the existential restriction) can be easily *checked in the Abox*, even complying to the open-world assumption (OWA). This favors the existential restriction over the universal restriction or max-cardinality restrictions, whose validation has to rely on the closed-world assumption (CWA).

OWL concept constructors' usability for focused categorization is analyzed in more detail in Section 5.

When considering the CE pattern $\exists P.C$, it may be practical to also consider some of its special cases or extensions that do not significantly worsen the tractability:

- One of them is $\exists P.\top$, the *unqualified existential restriction*, i.e., the restriction to the class `owl:Thing` in OWL terms, indicating that there is a property link but its filler does not matter when matching Abox data. The upper bound of the number of matching CEs is $|\mathcal{P}_O|$. (In Section 2.1 we already introduced the CEL $\mathcal{L}_{extop}$, which only allows this CE pattern.)
- The other is $\exists P.\{i\}$, i.e., the *individual value restriction*. The upper bound of the number of matching CEs is $|\mathcal{P}_O|.|\mathcal{I}_O|$, but $|\mathcal{I}_O|$ is rarely high, since ontologies normally contain few individuals; in addition, such 'internal codelist' individuals have their overarching class already assigned within the ontology, which may lead to pruning the number of relevant properties (only leaving those with the particular class in the range of the property).

Along with the CE patterns based on existential restriction, we will also consider the sole CE pattern of $\mathcal{L}_{nam}$ (the language of named classes) mentioned in Section 2, since we believe it should be by default a part of any practical CEL setting.

Based on these consideration (and, additionally, eliminating some uninteresting edge cases) we defined a suitable CEL, for the sake of the empirical research described further in this paper, as follows.

Definition 5 *The* simple existential CEL *is the CEL* $\mathcal{L}_{SE} = (\{C, \exists P.\top, \exists P.C, \exists P.\{i\}\}, \mathcal{R}_{ncp})$ *with* $\mathcal{R}_{ncp}$ *allowing for each signature* $\mathcal{S} = (\mathcal{C}, \mathcal{P}, \mathcal{I})$ *only those substitutions that assign a named class from* $\mathcal{C}$ *to the class variable C, and a named property from* $\mathcal{P}$ *to the property variable P.*

Therefore, the CE patterns considered for the sake of the presented research are: a named class; an unqualified property restriction (i.e., one with $\top$ as the filler); a property restriction with a named class as the filler; and, an individual value restriction (a property restriction with a singleton class as the filler).

Since $\mathcal{L}_{SE}$ does not allow for concept nesting, it guarantees *finiteness* of the set $\{Cat_i\}$ of categories for any O and FC. With respect to the three *OWL 2 profiles*, OWL 2 EL, OWL 2 QL, and OWL 2 RL, $\mathcal{L}_{SE}$, though simple, is only within the scope of OWL 2 EL, since the individual value restriction is only possible in this profile. Since the role of the CEL in our context is however not to form a new ontology on which reasoners would be systematically applied, but merely to provide a ground on which the FCP calculation could be performed, this expressiveness aspect does not present any serious problem.

Table 1 gives an overview of the CE patterns:

- The first column assigns the CE patterns a numbering local to the given CEL, for convenience.[13]
- The second column indicates the structure of the CE pattern itself in DL notation; the CE patterns substitutions are already restricted according to $R_{\mathcal{L}_{SE}}$: C stands for a named concept (class), P for a role (object property), and i for an individual.
- The third column indicates which variables from the CE structure are to be substituted by corresponding $sig(O)$ elements. The set of variables in

---

[13]We denote these specific CE patterns in bold face, to differentiate them from CE pattern meta-variables using the math font with subscript notation (c_i). We apply this distinction also to focused category patterns (**p1**, **p2**, . . . vs. p_i) in the next section.

Table 1

Summary of CE patterns in $\mathcal{L}_{SE}$

	Structure	Sign.	Abox	FCE pattern
c1	C	C	1	**p1**
c2	$\exists P.\top$	P	1	**p2**
c3	$\exists P.C$	P, C	2	**p3**
c4	$\exists P.\{i\}$	P, i	1	**p4**

c2 is a subset of those of **c3** and **c4**; therefore, if we consider one or more CEs for **c3** or **c4** with some property P then we should also consider the CE for **c2** with this P.

- The fourth column measures the length of the Abox path (as number of triples) connecting the individual j to be assigned to the CE with entities ('responsible' for the assignment) substituted for variables from the third column, in other words, the minimal size of the SPARQL graph pattern to be used for the CE instance detection in the Abox. The SPARQL pattern will rely on the *rdf:type* predicate (**c1**), on the specific predicate P (**c2**, **c4**), or on both (**c3**). For example, for **c3** the templated SPARQL graph pattern (where 'a' stands for *rdf:type* and the blank node [] for the existentially quantified individual of the restriction) is

$$\texttt{?j} \quad P \quad \texttt{[a } C \texttt{] .}$$

- The order of the CE patterns in the table reflects the increased complexity of their detection in the Tbox using FCE patterns designed for $\mathcal{L}_{SE}$. Those FCE patterns (presented next in Section 4), identified by their numbers, are in the fifth column.

Looking back at the competency questions of Ren et al. [19] referenced in Section 1, we see that while the archetypes #3 ("What type of [CE] is [I]?") and #4 ("Is the [CE1] [CE2]?") assume named subclasses of the FC, i.e. are covered by the CE pattern **c1**, archetype #1 ("Which [CE1] [OPE] [CE2]?") assumes categorization via an object property and a target class (or individual), i.e. is covered by the CE pattern **c3** (or **c4**).

4. Focused category extraction patterns for $\mathcal{L}_{SE}$

Having defined $\mathcal{L}_{SE}$ as a particular CEL, we can now proceed to formulating FCE patterns for it. We will then also briefly review the algorithmic approach for pattern-based CE generation, and outline the usage of patterns in the (optional) named class generation step. While these two aspects can be possibly generalized to

any CEL, we stick to their variant relevant in the $\mathcal{L}_{SE}$ setting, for simplicity.

4.1. Inventory for $\mathcal{L}_{SE}$ FCE pattern axiom templates

In principle, the FCE patterns could feature any OWL constructs within their axiom templates (*Tpl*). However, since our main target is to support the reuse of ontologies for linked datasets, it is natural to primarily build the patterns from the language primitives that are prominent in such ontologies (or, vocabularies). By querying the LOV portal SPARQL endpoint,[14] we found out that across all (778) vocabularies, the most frequent RDFS/OWL predicates – of those endowed with logical semantics – were `rdfs:subClassOf` (about 55K), `rdfs:domain`, and `rdfs:range` (about 32K each). Then followed `rdfs:subPropertyOf` (about 15K), the disjointness predicate `owl:disjointWith` (about 11K), the existential restriction predicate `owl:someValuesFrom`, and the universal restriction one `owl:allValuesFrom` (about 10K each[15]).

While the construction of the heuristic FCE patterns for $\mathcal{L}_{SE}$ – see the rest of this section – was not primarily guided by this frequency statistics but rather by long-term ontology designer experience, the inventory of most patterns indeed reflects the prominent role of `rdfs:subClassOf`, `rdfs:domain` and `rdfs:range` in linked data vocabularies.

Existential restriction is also used, but it is in the constraint part of the FCE patterns rather than in the axiom template part. Therefore, while the pattern setting assumes the presence of both 'global' (domain and range) and 'local' (existential) restrictions, omission of the latter by the designers does not preclude the generation of CEs but only disables their pruning.

4.2. FCE pattern collection for $\mathcal{L}_{SE}$

We present below five FCE patterns for $\mathcal{L}_{SE}$ and illustrate them on our running example of used cars.

---

[14] `https://lov.linkeddata.es/dataset/lov/sparql`, as of July 19, 2022

[15] Both the existential and universal restriction require three triples in the RDF representation; one of them in both cases relies on an auxiliary predicate `owl:onProperty`; its frequency (about 30K) is indicative of the total number of property restrictions of all kinds, incl. the cardinality ones.

Pattern **p1**   Let us first consider the FCE pattern for the CE pattern **c1** of $\mathcal{L}_{nam}$; the axiom templates are serialized as triples in the Turtle[16] RDF notation:

$$\mathbf{p1} = (\mathbf{c1}, (C), \{\, C \texttt{ rdfs:subClassOf } FC \;.\, \},$$
$$\emptyset, \mathbf{m1}) \quad (5)$$

where *FC* is the FC-variable and **m1** is a function returning its only input, the *named class* bound to the variable *C*. For example, if the FCP of an ontology *O* with respect to class `Car` is being computed, and the sole axiom template of pattern **p1** is instantiated by

`Truck rdfs:subClassOf Car .`

m1 will simply return the CE `Truck`, and the created category will thus be `Car` $\sqcap$ `Truck`. Note however, that in general the subclass relationship (to satisfy **p1**) can be both direct or indirect, and possibly even inferred using other kinds of axioms, i.e., the class matched to *C* can be any subclass of *FC* in the deductive closure of the ontology computed by a reasoner.

Also note that even such a simple FCE pattern might not be the only possible for the given CE pattern. If we accept the possibility that not all classes in different hierarchical paths are pairwise disjoint, we could also apply to **c1** a pattern for possibly unrelated *C* and *FC*, with the axiom template set being, e.g.,

$$\{\, C \texttt{ rdfs:subClassOf owl:Thing }.$$
$$FC \texttt{ rdfs:subClassOf owl:Thing } . \,\} \quad (6)$$

and with the constraints in *Cons* checking that none of $C \sqsubseteq FC$, $FC \sqsubseteq C$, $C \sqsubseteq \neg FC$ and $FC \sqsubseteq \neg C$ holds. Since the weight of the expressions built using such an FCE pattern should be rather tiny on average, we omitted it in our study, for simplicity.

As we move from $\mathcal{L}_{nam}$ to compound CEs of $\mathcal{L}_{SE}$, the FCE patterns are somewhat less obvious. We chose to implement them primarily in RDFS terms, namely, over `rdfs:subClassOf`, `rdf:type`, `rdfs:domain` and `rdfs:range` axioms, since these prevail in most linked vocabularies that are the prime subject in our ontology reuse setting. We however so far avoided `rdfs:subPropertyOf` to keep the pattern structure simpler (we will consider adding it in the future).

The relationship between `rdfs:domain` and the existential restriction will be central in our conception of FCE patterns for $\mathcal{L}_{SE}$. Note that for any *C*, *FC* and *P* holds that from the validity of *P* `rdfs:domain` *FC* we can trivially infer $\exists P.C \sqsubseteq FC$. The existential CE will thus necessarily be a subclass of the focus class *FC* in such a case; the only issue remaining is whether this subclass will have an 'interesting size' (neither too large nor too small) as subcategory of *FC*, in the spirit of the discussion from Section 2.3.

Pattern **p2**   The FCE pattern for CE pattern **c2** assumes that if the focus class *FC* is the domain of a property *P*, the instances of *FC* can be categorized according to whether they appear in the subject position in a triple with this property, whatever the object is (which truly corresponds to the existential restriction with filler $\top$, at the CE level). Again, even cases when *FC* is *inferred* as domain of *P* are considered.

$$\mathbf{p2} = (\mathbf{c2}, (P), \{\, P \texttt{ rdfs:domain } FC \;.\, \},$$
$$\mathbf{Cons2}, \mathbf{m2}) \quad (7)$$

Here, notably, the constraint set, **Cons2**, is not empty. Currently we only consider one constraint in **Cons2**: such that *P* must not appear in an existential restriction[17] $FC \sqsubseteq \exists P.C$; for such properties the CE would contain *all* instances of *FC*. As regards **m2**, it is, again trivially, a function that generates the existential restriction with filler $\top$ from its single input entity, the named property bound to the variable *P*.

In order to illustrate the application of the two FCE patterns introduced so far together, in the calculation of the approximate FCP (Eq. 4),we will return to our running example. Unlike in the previous, artificial, examples, we will refer to a real-world ontology.

Example 3 The *schema.org* ontology contains a class `Vehicle`[18] having four subclasses: `Car`, `BusOrCoach`, `Motorcycle` and `MotorizedBicycle`. It also proposes 82 properties having `Vehicle` in their domain.[19]

---

[16]https://www.w3.org/TR/turtle/

[17]The restriction can also be inherited from a superclass or part of a complete definition, or can have the form of a `value` or `self` restriction or of a cardinality restriction that entails the existential one; analogously for similar constraints for other FCE patterns.

[18]Here and in the rest of the running example we omit the `http://schema.org/` and other prefixes, for brevity.

[19]There is actually a tweak in this example. *schema.org* uses *schema:domainIncludes* instead of *rdfs:domain*, and analogously for

The instantiated $\widehat{FCP}$ formula (cf. (4)) for $\mathcal{CP}_{\mathcal{L}} = \{\mathbf{c1}, \mathbf{c2}\}$ (the respective $\mathcal{L}$ thus only covering a fraction of the expressiveness of $\mathcal{L}_{SE}$), $PS = \{\{\mathbf{p1}\}, \{\mathbf{p2}\}\}$, $FC = \texttt{Vehicle}$, $O = schema.org$, then becomes

$$\widehat{FCP}(FC, O, \mathcal{L}) =$$
$$Occ(\{\mathbf{p1}\}, FC, O) \cdot w^P(\mathbf{c1}) +$$
$$+ Occ(\{\mathbf{p2}\}, FC, O) \cdot w^P(\mathbf{c2}) =$$
$$4 \cdot w^P(\mathbf{c1}) + 82 \cdot w^P(\mathbf{c2}) \quad (8)$$

Intuitively, while $w^P(\mathbf{c1})$ should be close to the theoretical maximum (named subclasses are by default meaningful subcategories), $w^P(\mathbf{c2})$ should be much smaller. Positive examples of *schema.org* properties wrt. $\mathbf{p2}$ are $\texttt{review}$ (by substituting this property for P in $\exists P.\top$ we yield the category of cars for which at least one review exists), $\texttt{knownVehicleDamages}$ (category of cars that suffered from at least one damage) or $\texttt{vehicleSpecialUsage}$ (category of cars that have been used in a special mode, e.g., as a cab). Properties such as $\texttt{fuelType}$ or $\texttt{numberOfDoors}$ would, in turn, only yield meaningful categories when filled with a certain value – since *all* cars are supposed to consume *some* fuel (or another source of energy) and have *some* number of doors. The inclusion of their $\mathbf{p2}$-matches in the computation will thus make $\widehat{FCP}$ slightly overstated.

Let the weights be, for example, $w^P(\mathbf{c1}) = 1$ and $w^P(\mathbf{c2}) = 0.25$. This would yield $\widehat{FCP} = 4 \cdot 1 + 82 \cdot 0.25 = 24.5$. The categorization power of *schema.org* with respect to the focus class $\texttt{Vehicle}$ would thus be for example greater than that of a purely taxonomic ontology in which a class (semantically corresponding to) $\texttt{Vehicle}$ would have 20 transitive subclasses, but smaller than that of a purely taxonomic ontology with 30 such subclasses.

Let us now proceed to the FCE patterns for the remaining CE patterns included in $\mathcal{L}_{SE}$.

---

the range axioms (*schema:rangeIncludes*). These properties do not have inferential semantics, and merely relate a given property to a class that "is (one of) the type(s) the property is expected to be used on" (`https://schema.org/domainIncludes`), and "constitutes (one of) the expected type(s) for values of the property" (`https://schema.org/rangeIncludes`), respectively. However, if a property only has a single *schema:domainIncludes* axiom, which is the case for most properties linked to *schema:Vehicle*, the heuristic use of the pattern is similarly plausible as for *rdfs:domain*.

Pattern $\mathbf{p3}$ The CE pattern $\mathbf{c3}$ includes all instances of FC that appear in the subject of a triple with P as predicate and some i as object such that i is an instance of class C. The FCE pattern for $\mathbf{c3}$ then extends $\mathbf{p2}$ such that we also consider the range of P, which is an auxiliary class D. Since D itself would not make a proper distinction (all instances satisfying $\mathbf{c2}$ will by the definition of the $\texttt{rdfs:range}$ predicate be connected by P to an instance of D), we further need to involve a 'refining' class C as subclass of D, which is also reflected by the second mapping variable.

$$\mathbf{p3} = (\mathbf{c3}, (P, C), \{\, P \text{ } \texttt{rdfs:domain } FC \text{ ;}$$
$$\texttt{rdfs:range}_a \text{ } D \text{ . } C \text{ } \texttt{rdfs:subClassOf } D \text{ . } \},$$
$$\mathbf{Cons3}, \mathbf{m3}) \quad (9)$$

The inferential closure is again used, however, with the exception of the range axiom, which is only considered as asserted (therefore the 'a' index in $\texttt{rdfs:range}_a$) – otherwise not only subclasses of D but also classes having a common superclass with D would be returned as C (since the superclass is also inferred as a range of P). The constraint in $\mathbf{Cons3}$ states that P must not appear (in the inferential closure of O) in an existential restriction $FC \sqsubseteq \exists P.E$ such that $E \sqsubseteq C$; for such properties the CE would contain *all* instances of FC. The mapping function $\mathbf{m3}$ is, again, trivial, except that its input is now a pair (P, C), from which the qualified existential restriction $\exists P.C$ is built.

As this FCE pattern is already a bit more complex, we will also demonstrate its application (and the added value compared to $\mathbf{c2}$) graphically, in Fig. 1. The upper part of the diagram displays the structure (Tbox) of $\mathbf{p3}$; the bold labels indicate asserted statements, and the italic label indicates an inferred statement (here we show a case where FC only belongs to the domain of P by 'inheritance' from its subclass). The bottom part shows example constellations of an instance and non-instance of $\exists P.C$ in the Abox. While both x_1 and x_2 belong, via inference over the domain axiom, to FC, and both also belong to $\exists P.\top$ and (via the inference over the range axiom) to $\exists P.D$, it is only x_1 that also belongs to $\exists P.C$ as the considered category.

Reliance of FCE patterns on *asserted* ontology structures might make the pattern vulnerable to ontology versions that would already contain the inferences materialized. However, while the transitive *rdfs:subClassOf* inference is sometimes materialized
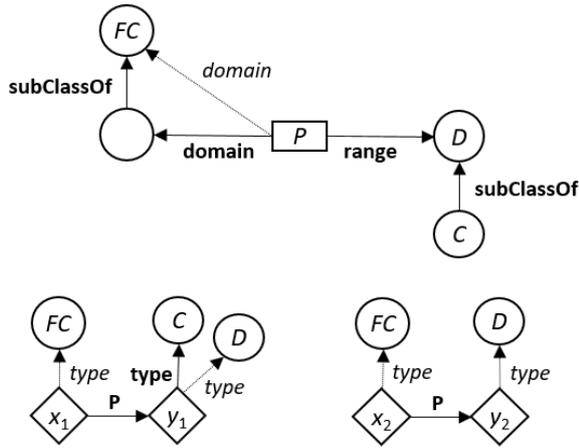
Fig. 1. Tbox of **p3** and Abox of instance and non-instance of $\exists P.C$

in ontologies (transitive superclasses being present for a class, rather than just its direct superclass), are linked to it in the 'asserted' ontology), it is very unlikely that this would be the case for the upward propagation of *domain/range restrictions*.

Pattern **p4** The CE pattern **c4** includes all instances of *FC* that appear in the subject of a triple with *P* as predicate and (the specific individual) *i* as object. The FCE pattern for **c4** then extends **p3** with an individual that is instance of *C* and is also used in the mapping variable sequence (replacing there the class *C*):

$$\mathbf{p4} = (\mathbf{c4}, (P, i), \{ \; P \; \texttt{rdfs:domain} \; FC \; ;$$
$$\texttt{rdfs:range}_a \; D \; . \; C \; \texttt{rdfs:subClassOf} \; D \; .$$
$$i \; \texttt{rdf:type} \; C \; . \; \}, \mathbf{Cons4}, \mathbf{m4}) \quad (10)$$

The inferential closure is used as in **p3**, i.e., with the exception of the range axiom. The constraint in **Cons4** states that there must not be (in the inferential closure of *O*) an existential restriction $FC \sqsubseteq \exists P.\{i\}$; for such properties the CE would contain *all* instances of *FC*. The mapping function **m4** is analogous to **m3**, except it wraps the individual bound to *i* in a singleton enumeration.

Pattern **p5***: SKOS concept as value* For the CE pattern **c4** we include one more FCE pattern, which goes beyond the RDFS expressiveness and relies on the connection of the ontology to an external resource – a *SKOS*[20] *codelist*. This variant thus extends the previ-

---

[20] http://www.w3.org/2008/05/skos

ous one for a specific source of instance *i*, which is a code list (also called concept scheme in SKOS):

$$\mathbf{p5} = (\mathbf{c4}, (P, i), \{ \; P \; \texttt{rdfs:domain} \; FC \; ;$$
$$\texttt{rdfs:range}_a \; \texttt{skos:Concept} \; ,$$
$$[\; \texttt{a} \; \texttt{owl:Restriction} \; ; \; \texttt{owl:hasValue} \; s \; ;$$
$$\texttt{owl:onProperty} \; \texttt{skos:inScheme} \;] \; .$$
$$i \; \texttt{a} \; \texttt{skos:Concept} \; ; \; \texttt{skos:inScheme} \; s \; . \; \}$$
$$\mathbf{Cons4}, \mathbf{m4}) \quad (11)$$

The difference is merely in the selection method for *i* – rather than an instance of a class from the current ontology, it has to be a `skos:Concept` linked to a concept scheme *s* that is in the range of *P*. The inferential closure is used as before, and the constraint and the mapping function are the same as in **p4**.

We can now illustrate the application of all considered FCE patterns on our used cars example.

Example 4 The property `vehicleSpecialUsage` of *schema.org* has the class `CarUsageType` as its range. `CarUsageType`, in turn, has no subclass but three enumerated instances: `DrivingSchoolVehicleUsage`, `RentalVehicleUsage` and `TaxiVehicleUsage`. It would thus yield three categories via pattern **p4** (and none for **p3**), in addition to the one for **p2** from Example 3. In contrast, the properties `fuelType` or `numberOfDoors` only have a disjunction of generic classes (`QualitativeValue`, `Text` and `URL`) in their range.[21] They therefore do not yield any category via **p4**, within *schema.org*. However, if the range of such properties were later associated with relevant SKOS code lists containing fuel types and possible numbers of doors, respectively, **p5** could be applied.

All five FCE patterns assure (**p2**–**p5** via the domain axiom) that the mapped CE is a *specialization* of the FC.[22] Therefore, any $\mathcal{L}$-category $Cat = FC \sqcap D$ generated from an ontology using these patterns is equivalent to its differentiating class expression *D*. It is also easy to see that the patterns, except **p4** vs. **p5**, are *mutually exclusive* since they produce structurally differ-

---

[21] Such special-purpose classes would have to be filtered out; typically they could be automatically detected by appearing in the range of a huge proportion of properties.

[22] In contrast, this would for example not be the case for the alternative FCE pattern for **c1** (cf. Equation 6) allowing hierarchically unrelated *C* and *FC*.

ent CEs. (Formally, to assure the exclusivity of **p2** and **p3**, *D* in **p3** should not be `owl:Thing`. We however do not anticipate that explicit range axioms would have the default value `Thing`.)

As regards the complexity of the FCE patterns in terms of (templated) RDF triples in the template axioms, it is as listed, for the respective CE patterns, in the fifth column of Table 1. **p4** is the most complex one of the 'straightforward' FCE patterns, with four triples; however, **p5** (for the same CE pattern, **c4**) is even more complex, having eight triples overall, due to the overheads of representing OWL restrictions in RDF.

We have previously stated that the proposed FCE patterns are merely heuristic (approximate) with respect to the optimally chosen set of $\mathcal{L}_{SE}$ categories. Let us now briefly elaborate on this statement and summarize its different aspects:

- First of all, the Assumption #4 from Section 1, i.e., that the quality of compound categories should be correlated with the degree to which users would consider a corresponding named class as meaningful, rules out any kind of optimality guarantee, since human judgment is always subjective. The designer wishing to reuse an ontology may even question the *asserted subclass axioms* from the source ontology, thus lowering the *precision* of the **p1**. The (human-judged) precision of FCE patterns for *compound CEs* will naturally be even less stable.
- The situation is analogous for the FCE patterns' *recall*. Even **p1** will omit any category with a differentiating class expression (DCE) beyond those subsumed by the FC. Typical cases of 'lost' categories are those with the FC and the DCE being 'close neighbors' having a common 'natural' superclass (e.g., *Student* and *Teacher* being subsumed by *Person*).
- As regards the FCE patterns **p2**–**p5**, which rely on domain/range axioms in their axiom templates and on (negated) existential restrictions in their validity constraints, their accuracy depends on the degree in which the ontology is equipped by these kinds of axioms; see Section 8.2 for a more detailed discussion.

4.3. Named category generation

The generic algorithm in Listing 1 ends by generating a list of CEs (plus their CE patterns) for a given fo-

cus class. However, to finalize our *schema.org* example narrative, we will also consider the subsequent step of *named category generation*, already briefly mentioned in Example 1.

Example 5 Let us assume that the categories discovered for the purpose of the FCP computation are retained for future use. The vehicle retailer might then decide to rebuild the product catalog so as to cover the categories, even the compound ones, that are populated by a significant number of vehicle items. The categories derived from the same property, say, `vehicleSpecialUsage`, would then generate a branch of the catalog structure, which, after some (possibly, NLP-assisted) tweaks could look like this:

```
Vehicles
   Vehicles with special usage
      Vehicles with driving school usage
      Vehicles with rental usage
      Vehicles with taxi usage
```

This structure might either become materialized in the underlying ontology as such or might only be generated at the web engineering level. The taxonomy would naturally follow the specialization of CE patterns as in Table 1: the categories with DCE conforming to **c2** would be direct subcategories of the FC (and siblings of those with DCE conforming to **c1**, i.e. of the original named subclasses); the categories with DCE conforming to **c3** would be subcategories of that with DCE conforming to **c2** (for the same property *P*), and the categories with DCE conforming to **c4** would be either at the same level as those of **c3**, or some might be a subcategory of a category with DCE conforming to **c3** if there were an `rdf:type` link between the instance *i* in the match of **p4** and a subclass of the class in the range of *P* in the match of **p4**. (Since the derivation of named categories is not the main topic of this paper, we omit the formalization of this derivation for space reasons, however simple it is.)

4.4. Summary

FCE patterns are a crucial element of the operationalization of FCP computation over the ontology axioms. We therefore described the FCE patterns for $\mathcal{L}_{SE}$ in detail in subsection 4.2, and discussed the rationale of their construction. As it can be observed, subsection 4.3 primarily refers to CE patterns rather than to their associated FCE patterns; it was mainly included here due to Example 5 being a close sequel to the running example in subsection 4.2 (Example 4).

5. CE patterns in ontology axioms

We analyzed the axioms of publicly available ontologies as an empirical source for estimating the frequency of occurrence of CE patterns of the CEL $\mathcal{L}_{SE}$

5.1. Ontology axiom sources

For our analysis we used three collections of ontologies:

1. The collection indexed by the *Linked Open Vocabularies* (LOV) portal,[23] containing real-world (but mostly light-weight) ontologies typically used for linked data publishing.
2. The BioPortal collection,[24] the well-known repository of biomedical ontologies.
3. A small experimental collection of ontologies having heterogeneous styles and relatively rich in axioms, from the domain of conference organization, called *OntoFarm*.[25]

The analysis took place in November 2017.

5.2. Motivation of the analysis

The impetus for this empirical analysis was the close relationship between the central motivating task of the research, that of using CEs (either named classes, or compound CEs that can possibly be transformed to named classes) for *categorizing individuals*, on the one hand, and the task of *defining new named*[26] *classes* (using compound CEs) within ontologies on the other hand. We conjecture that the distribution of compound CE structures in both tasks would be to some degree similar. The differences in the relevance of constructs wrt. these tasks are discussed below.

Note that the subsequent use of the constructed CEs is, in some aspects, similar in both tasks, too. Notably, in all cases, an aspect of *categorization* is present:

– A compound CE in the RHS of an *equivalence axiom* allows to *intrinsically* (i.e. using the apparatus of DL tightly coupled with OWL as rep-

resentation language) *infer* individuals to be instances of the named class in the left-hand side (LHS) of the axiom.
– A compound CE in the RHS of a *subsumption axiom* (which represents a mere necessary, and not sufficient, condition) allows, analogously, to *intrinsically rule out* individuals from being instances of the named class in the LHS of the axiom.
– A compound CE that can be merely constructed from the *ontology signature* under some CEL and *for some FC* still allows to subcategorize individuals, but merely *extrinsically*, i.e., either using manual assignment or some other external source (e.g., a machine-learning-based classifier).

Technically, the expected outcomes of the analysis were the following:

– Findings about the frequency of various concept expression patterns in the RHS of axioms in existing ontologies (possibly interesting for the community even beyond the focused categorization setting)
– Positioning of the CEL $\mathcal{L}_{SE}$ with respect to the rank of its CE patterns in a CE pattern frequency table.

The considered list of CE patterns is, essentially, that of first-level constructors in the axiom RHS. However, since we were particularly interested in the compound CE patterns of $\mathcal{L}_{SE}$, we singled them out under the labels 'ExistThing' (**c2**), 'ExistNamed' (**c3**) and 'ExistIndiv' (**c4**). All other variants of existential restrictions not falling under these three are counted under 'ExistAnon', since the filler then is a compound CE rather than a named class.[27]

The frequency of CE pattern occurrence, in absolute counts, was classified by three dimensions:

– By the analyzed ontology collection.
– By the distinction of equivalence or subclass axioms (and the sum of both).
– By the level of nesting: outermost constructor vs. further levels.

---

[23]https://lov.linkeddata.es/dataset/lov/

[24]https://bioportal.bioontology.org/

[25]https://owl.vse.cz/ontofarm/; the collection is used, among other, in the Ontology Alignment Evaluation Initiative, http://oaei.ontologymatching.org/.

[26]We can for now ignore general concept inclusions (having a compound CE as their left-hand side), which are allowed in some dialects of OWL but only used sparingly in real-world ontologies.

[27]Strictly speaking, the singleton enumeration $\{i\}$ in **c4** is also an anomymous CE, and the notion of 'individual value restriction' is rather a syntactic sugar in DL terms. For our purposes it is however distinguished from the restrictions that have some other compound filler.

Table 2

CE patterns in axiom RHS

LOV			
Equivalence		Subclass	
Construct	#	Construct	#
Conjunction	1134	*ExistNamed*	3187
ExistNamed	126	Universal	1331
Disjunction	108	Conjunction	553
Enumeration	54	Exactly	545
ExistAnon	31	Minimally	483
ExistIndiv	30	ExistAnon	467
Exactly	12	Maximally	237
Negation	4	*ExistIndiv*	77
Minimally	3	Disjunction	56
Maximally	1	Negation	23

BioPortal			
Equivalence		Subclass	
Construct	#	Construct	#
Conjunction	122725	*ExistNamed*	1977805
ExistAnon	24239	ExistAnon	64046
ExistNamed	2338	Conjunction	18409
Enumeration	854	Universal	6587
Disjunction	658	*ExistIndiv*	4698
Universal	22	Exactly	1179
ExistThing	11	Disjunction	856
Minimally	10	Minimally	579
Exactly	7	Maximally	577
Negation	6	Negation	96

OntoFarm			
Equivalence		Subclass	
Construct	#	Construct	#
ExistNamed	40	*ExistNamed*	317
Conjunction	27	Universal	47
Disjunction	15	Exactly	23
Universal	2	Minimally	22
Enumeration	1	ExistAnon	15

5.3. Results and their discussion

The results are in Table 2, however omitting the last dimension (nesting level) for brevity. The most frequent constructors are listed: 10 for LOV and BioPortal, and 5 for OntoFarm (where only the top of the ranking is relevant, due to very low counts). The constructors corresponding to $\mathcal{L}_{SE}$ are in italics.

As regards the $\mathcal{L}_{SE}$ patterns, we can see that 'ExistNamed' ($c3$) appears very often at top positions, throughout the collections and axiom types; it dominates in subclass axioms, most notably in BioPortal. 'ExistIndiv' ($c4$, i.e. the value restriction) is less fre-

quent but still plays some role in both larger collections. 'ExistThing' ($c2$) appears even less often.

The other constructors appearing at the first level of the axiom RHS nesting are:

– *Conjunction* ($\sqcap$). The possibility of building logical conjunctions via plausible axiom patterns, analogous to those from Section 4, deserves a further study. E.g., for the `Car` focus class, *conjunctions of property restrictions* such as

$$\exists \texttt{hasSeller.Corporate}$$
$$\sqcap \; \exists \texttt{usage.TaxiUsage}$$

(where the two properties are required to share their domain) could provide further weight to an ontology in the FCP computation, atop the weight of their constituent property restrictions. This would however lead to a significantly more complex CEL. In contrast, considering mere *conjunctions of named classes* (which would still fit into a 'simple' CEL) would be problematic due to the combinatorial explosion of class pairs, in which only a tiny fraction would be mutually compatible (considering that the absence of explicit disjointness axioms does not guarantee class compatibility in real-world ontologies). Also note that from the point of view of data description and querying as such, emulating conjunctive concepts is simple: multiple categories are just applied at the same time, either to categorize a given individual or to form a simple query (e.g., by ticking multiple categories in the product catalog form). With respect to the above example, the two newly constructed named classes, as categories assigned/ticked in combination, might be called `CarWithCorporationSeller` and `CarWithTaxiUsage`, respectively.

– *Disjunction* ($\sqcup$). Unlike for conjunctive expressions, any pair of expressions is mutually compatible in a disjunction. As mentioned in Section 3, the number of disjunctive combinations will thus be very high for larger ontologies, and the question of which combinations should be taken into account for FCP or not would not be easy to decide at the axiom pattern level. Intuitively, concepts 'closer to each other' in the ontology might be more adequately combined (e.g., the customer might be more likely to seek a "`Sedan or Hatchback`" than a "`Bus or Bicycle`)", and

similarly in combining property expressions for properties with the same domain. However, such a heuristic would favour shallow ontologies over ontologies with a more elaborate class taxonomy. The utility of disjunctively defined categories in RDF data publishing practice (e.g., within e-shop catalogs) deserves a further study.

– *Existential restriction to a compound CE* ('ExistAnon'). This could be one of the meaningful extensions to $\mathcal{L}_{SE}$. Instead of only allowing a named class as property filler (as in **c3**), we would allow a certain subset of compound CEs. The weight of such expressions in the FCP computation would of course have to be adequately decreased compared to less complex expressions. We will study this option in our future research; one promising candidate CE pattern would be the embedding of two existential restrictions (corresponding of a two-piece property chain).

– *Universal restriction* ($\forall$). The observation that universal restrictions rarely appear in equivalence axioms can be explained by the fact that classifying individuals by universal restrictions is impossible on the open semantic web due to the OWA: the fact that all *known* individuals y_i such that $x\ P\ y_i$ belong to C does not guarantee that $x \in \forall P.C$ holds, because there may exist some *unknown* y_j such that $x\ P\ y_j, \quad y_j \neg \in C$. Yet, universally restricted concepts may be valid categories in the FCP context, since the usage of the ontology involving the 'extrinsic' categorization may already be performed under the CWA. For example, the used vehicle retailer may have trustful evidence that *all* previous owners of the car have been corporations. In terms of heuristically deriving a universal restriction category (with a named class as filler) from the Tbox, there would be probably very little difference from the existential restriction (to the same named class): the same Equation 9 (based on a domain and range axiom) could be applied, except that the pattern constraint should be changed to a universal restriction (ruling out the possibility that the category would coincide with the focus class). Additionally, we would assume that a universal property restriction category would be implicitly applied in conjunction with the existential one. For example, based solely on the universal restriction, the CE $\forall$hadAccident.FatalAccident would include all cars not having any accident recorded in their history at all. To avoid this ef-

fect, we should extend it to $\exists$hadAccident.$\top$ $\sqcap$ $\forall$hadAccident.FatalAccident.

– *Cardinality restrictions* ($\leq_n$, $\geq_n$, $=_n$, i.e., 'Minimally', 'Maximally' and 'Exactly' in Table 2). They appear relatively frequently in subclass restrictions, much less so in equivalence ones. While some of them may be valid categories in the FCP context, their heuristic derivation from the Tbox would require to specify the numerical parameter. As in the case of the universal restriction, the heuristic pattern for the generation of cardinality restrictions would basically be the same as for the existential restriction (Equation 9), only the pattern constraints would differ.

– *Enumeration* ($\{\dots\}$). Using enumerations in practical focused categorization scenarios is mostly irrelevant. Enumerations presumably appear as the outermost level in the RHS of (equivalence) axioms whose LHS is a class serving as an embedded (thus typically small), intentionally closed 'code list' inside the ontology. The 'code list class' is then not used on its own, but rather helps define new concepts. The broadly usable 'code' individuals embedded in the ontology are then unlikely to coincide with the ad hoc instances of a focus class that we would like to subcategorize.

– *Negation* ($\neg$). In some contexts, a negated concept might perhaps be a sensible subcategory of a focus class; however, such contexts would be hard to guess a priori. Considering the negation categories in a generic way would simply double the number of categories for each CE pattern to be negated, and thus would not have other impact on the FCP than would have the doubling of the weight of those CE patterns.

5.4. Summary

The contribution of this section was twofold. First, it featured the CE pattern occurrence frequencies as such, which could be of general interest to the ontology design community. Second, specifically in the line of the current paper, it provided intellectual analysis of axiom type roles in FCP.

The analysis suggested that the $\mathcal{L}_{SE}$ patterns play a significant role in the family of all compound expressions commonly used in OWL. The design of an FCP formula restricted to this CEL is thus meaningful, considering that we want to investigate the landscape of the whole focused categorization problem within a simple language first. The analysis however also indi-

cated directions for extending the used language in the future – most notably, in relation to the expected complexity of the *queries* to be applied on data published using the reused ontologies.

6. FCE pattern occurrence in the Tbox of ontologies

The questions to be answered by this analysis were:

1. How many ontologies, and for how many FCs, provide a decent number of 'categorizing' CEs through heuristic mapping from the patterns from Section 4.
2. What are the differences in the occurrence of the individual FCE patterns overall and across different ontology collections.

The answers to these questions would help us estimate how likely it is that the particular FCE pattern would yield potential categories when a focus class is provided as an input. (They will not, though, provide an estimate of the quality of such categories; this question will be addressed in Section 7.)

6.1. Ontology sources and data aggregation method

In the analysis we made use of our *Online Ontology Set Picker framework*[28] (a tool allowing to quickly select ontologies according to various metrics) to process ontologies from two collections. One is LOV; we again used its November 2017 snapshot where 617 ontologies were available, of which we could process successfully 568 at syntactical level. The other is Onto-Farm, which contains 16 ontologies. We omitted Bioportal in this analysis, since it is on average rather poor in properties, which are a necessary ingredient for generating categories within $\mathcal{L}_{SE}$.

In order to provide aggregate results, we counted the occurrences of FCE patterns from Section 4 across all classes of all ontologies evaluated in the role of FC. We summed up these results at ontology level by identifying 'categorizable' classes, i.e. classes for which the pattern occurrence reached some threshold τ (1, 3, and 5) for patterns **p1**, ..., **p4**;[29] see Table 3. The tables show the percentage of ontologies for which more than a given portion of classes, *ctg*, is 'categorizable'

Table 3

Ratio of ontologies with a > c portion of classes 'categorizable'

OntoFarm						
Pat.	τ	> 0.0	> 0.1	> 0.2	> 0.3	> 0.4
p1	1	100.0%	100.0%	81.3%	25.0%	0.0%
	3	100.0%	100.0%	18.8%	0.0%	0.0%
	5	93.8%	31.3%	0.0%	0.0%	0.0%
p2	1	100.0%	93.8%	56.3%	31.3%	25.0%
	3	100.0%	43.8%	25.0%	12.5%	0.0%
	5	93.8%	37.5%	6.3%	0.0%	0.0%
p3	1	93.8%	50.0%	18.8%	6.3%	6.3%
	3	93.8%	50.0%	12.5%	0.0%	0.0%
	5	93.8%	31.3%	**6.3%**	0.0%	0.0%
p4	1	25.0%	0.0%	0.0%	0.0%	0.0%
	3	18.8%	0.0%	0.0%	0.0%	0.0%
	5	6.3%	0.0%	0.0%	0.0%	0.0%
LOV						
Pat.	τ	> 0.0	> 0.1	> 0.2	> 0.3	> 0.4
p1	1	80.1%	73.1%	47.7%	25.4%	11.6%
	3	60.4%	43.0%	13.0%	3.2%	1.4%
	5	49.3%	22.4%	3.2%	1.4%	1.1%
p2	1	83.8%	72.4%	60.7%	46.8%	34.5%
	3	71.1%	51.6%	29.8%	16.2%	8.6%
	5	56.2%	32.7%	14.4%	7.2%	3.9%
p3	1	49.3%	29.0%	14.1%	6.2%	3.0%
	3	42.1%	20.6%	8.8%	3.2%	**1.8%**
	5	34.2%	14.1%	**5.5%**	**1.9%**	**1.4%**
p4	1	22.0%	9.9%	4.6%	2.1%	1.2%
	3	18.1%	7.2%	3.3%	1.2%	0.9%
	5	16.2%	6.9%	3.2%	1.2%	0.9%

as FCs using the given pattern. The portion of 'categorizable' classes is quantized by 0.1, from >0.0 to >0.4; higher values are not so relevant, since, intuitively, not all classes of an ontology but only some proportion of more general ones would likely serve as FCs.

For illustration of the bottom-up calculation steps, let us take the example of an OntoFarm ontology called *cmt*. This ontology satisfies the parameter thresholds for $\tau = 5$, $ctg > 0.1$, for **p1**, as 4 of its 29 classes ($n \sim 14\% > 0.1$) are 'categorizable': each has five or more named subclasses (i.e. its τ is at least 5) . As we see in the respective cell of Table 3, *cmt* thus belongs to the 31.3% of OntoFarm ontologies having such a proportion of 'categorizable' FCs.

6.2. Results for patterns **p1** to **p4**

Let us now try to answer the questions posed at the start of the section, by examining the result table.

Unsurprisingly, the percentage of 'categorizable' classes is in most cases highest for **p1** and **p2**, which reflects the simplicity of these patterns: many classes have some named subclass (**p1**), and many also appear in the domain of some property (**p2**). With increasing *ctg*, the percentage drops more slowly for **p2** than for **p1**, reflecting the fact that even classes lowest in the subclass hierarchy (i.e. not having named subclasses) do appear in the domain of properties.

For **p3** the percentage is generally lower; however, it 'beats' **p1** when it comes to the setting with higher thresholds for both *ctg* and τ, i.e., when we want to include even less general (focus) classes and require them to have a higher number of categories. This is apparent even in the much smaller OntoFarm (for $ctg > 0.2, \tau = 5$), but particularly in LOV; we indicate such values using boldface in the table. Presumably, focus classes down in the hierarchy may be connected (through some properties' domain and range) to more general classes, thus 'borrowing' their subclass hierarchy (note the structure of **p3** in Equation 9, where class *C* from the range of property *P* 'lends' its subclass *D* for the filler of the existential restriction category).

p4 is least frequent, however, for higher *ctg* and $\tau = 3$ the percentage also gets closer to that of **p1** (presumably, through the same mechanism as we discussed for **p3**).

6.3. Results for pattern **p5**

For completeness, we performed an analysis of ontologies which use SKOS concepts for entity categorization (pattern **p5**). While OntoFarm ontologies do not feature this pattern at all, in LOV there were 7 ontologies out of 509 containing its instantiation, to date.

Table 4 presents information about those 7 ontologies in terms of the number of SKOS concepts that can be used as the 'categorization individual' and the number of SKOS concept schemes from which those concepts come. Further, we include information about the date of the last modification of the ontology. In two cases, no concept schemes are available. For the other ontologies the number of SKOS concepts (and SKOS schemes, respectively) usable for categorization varied from 11 to 274 (from 1 to 16, respectively). Although the phenomenon captured by **p5** (reference to a SKOS code list in an ontology) is currently marginal, we assume that it may have an increasing trend, since those seven ontologies had been last modified, on average, in 2014, while the average of the last update dates for the whole of LOV was in 2012.

6.4. Summary

The findings about **p3** and **p4** are interesting and potentially relevant for the efficiency of the FCP computation. If we managed to prove that the categories generated by these patterns are often meaningful (which will indeed be hinted by the cognitive experiments in Section 7), these two patterns could yield useful categories even for focus classes positioned lower in the subclass hierarchy, for which **p1** does not work well.

7. Cognitive experiments: CE assessment

The previous analyses carried out over the ontology Tbox (axioms RHS in Section 5 and FCE pattern occurrence in Section 6) only indirectly contributed to the central question: whether the compound CEs from $\mathcal{L}_{SE}$, and to what degree, can serve as meaningful subcategories of their focus classes.

In order to get finer insights, we proceeded to a detailed investigation of sample CEs by human 'ontologists', both experts and relative novices (students of relevant subjects). We performed two campaigns of experiments, the first in Spring 2016[30] and the second, with a slightly modified and extended setup, in Spring 2018. The overall scheme of the experiments was as follows:

1. Provide the human assessors with a set of 'focus class – subcategory' pairs, such that the subcategories correspond to *different CE patterns*. Each pair constituted an elementary *assessment task*.
2. Collect the *assessment* for each task, reflecting the degree to which the assessor perceived the category as 'meaningful', or, better, 'meaningful and reusable' (or also, for short, what is the perceived quality of the category), via a *questionnaire*.
3. *Aggregate* the collected data across the assessors.
4. Seek *correlations* between the category quality and the CE pattern, possibly corrected through *metadata* (also collected in the questionnaire) such as the level of written English of the assessors or their general comprehension of the meaning of the constituent entities of the category.
5. Examine the *qualitative feedback* also collected through the questionnaire.

---

[30]Already described in our early publication [23]. Here we provide a synoptic view of both campaigns.

Table 4

Categorization via SKOS concepts (FCE pattern **p5**)

Ontology	# of schemes	# of concepts	Last modified
http://www.loc.gov/premis/rdf/v1	16	79	2012-09-14
http://purl.org/procurement/public-contracts	4	86	2012-10-10
http://data.ign.fr/def/geofla	1	NA	2015-01-11
http://datos.gob.es/def/sector-publico/organizacion#	3	NA	2015-02-19
http://data.ign.fr/def/topo	13	274	2014-04-12
http://contsem.unizar.es/def/sector-publico/pproc	3	11	2015-01-11
http://rdf.insee.fr/def/geo	3	30	2015-01-11

Table 5

Overview of cognitive experiments with students

Campaign	Summer 2016	Summer 2018
# of students	27	31
# of different tasks	59	40
# of tasks/person	20	8
Avg. time/task	90 sec	5 min
Tasks source	LOV, OntoFarm	LOV
CE patterns	**c2**, **c3**, **c4**	**c1**, **c2**, **c3**, **c4**
Judgment support	Glossary of conference terms	Entity definitions from the ontologies; CE structure verbalization
Collected metadata and other info	Comprehension of entity semantics	English skills; comprehension of entity semantics; justification for negative assessment

A summary of the experimental setting in both campaigns is in Table 5.

7.1. Summer 2016 campaign

Initial sampling As regards the FC+subcategory (from now on, 'task') sampling for both threads of analysis (expert/novice), we used the same collections as in the survey from Section 6, i.e. LOV (a 2016 snapshot) and OntoFarm (this collection has remained unaltered for nearly a decade). From each collection 40 tasks were randomly generated while maintaining an approximately even proportion of subcategories for patterns **c2**, **c3** and **c4**, eventually yielding 80 tasks. After a manual removal of duplicates (for OntoFarm as smaller collection) and tasks involving entities with cryptic names without meaning in natural language, 59 tasks remained (28 from OntoFarm and 31 from LOV); there were 17 tasks for **c2**, 20 for **c3** and 22 for **c4**.

Expert ontologist assessment and insights. In the first stage, the assessment was made by three researchers with 10–20 years of experience in ontology

engineering (authors of this paper: VS, OZ and MV). They first examined the 59 tasks independently and assessed them on the 5-point Likert scale: for each task, with a focus class *Y* and a subcategory *X*, the question "Is *X* a meaningful subcategory of *Y*?" was answered as either 'certainly', 'perhaps', 'borderline', 'perhaps not' or 'certainly not'. Then a consensus was sought in a F2F session. The independent assessment had 76% agreement: in 45 out of 59 cases there was no contradictory assessment (certainly/perhaps yes vs. certainly/perhaps not); we will call these cases *clear positives* (42 cases) and *clear negatives* (3 cases), respectively. The consensus session then yielded a complete consensus on the remaining cases; in 12 out of the 14 'clash' cases the final result was 'yes' (namely, a conceivable situation was formulated in which the CE would be a meaningful subcategory of the FC), one case was found dubious due to an implausible inference (see the second 'insight' below, on the 'village head chef') and in one case the subcategory was assumed semantically equivalent to its FC, both resulting into 'no'. Of the five ultimately negative results, four were of **c2** and one of **c3**. Complete assessments and commentaries are in a supplementary page. Selected general insights into less obvious decisions, with examples, follow:[31]

- Ontologies tied to software applications, such as some OntoFarm ones (capturing the processes supported by conference software) use object properties to capture relationships that are only relevant within a *short time frame*, e.g., `cmt:finalizePaperAssignment`; a meaningful category of persons would rather refer to their long-term responsibility for paper assign-

---

[31]In this section we write the example CEs in Manchester syntax, with the keyword `some` indicating an existential restriction in **c2** and **c3** CEs and the keyword `value` indicating a value restriction in **c4** CEs.

ment than to the instantaneous action of 'finaliz-ing' it.

– In some cases the use of inferential closure for the filler class in **c3** leads to linking relatively *thematically unrelated* entities (especially in the DBpedia ontology), such as in `dbo:beatifiedPlace some dbo:WineRegion` for instances of the FC `dbo:Person`. This case was still found marginally acceptable, as there could be some correlation between religiosity and wine production. However, a similar one, with subcategory `dbo:headChef some dbo:BaseballPlayer` proposed for the FC `dbo:Village`, was rejected not only due to the thematic leap but also due to an odd inference result: the declared domain of `dbo:headChef` is `dbo:Restaurant`, but the ontology (actually, the 2014 version from the LOV endpoint) enabled to infer the axiom `dbo:Restaurant rdfs:subClassOf dbo:Village`.

– Some CEs of **c4** are plausible but less useful due to their *inherently limited extent*: for instance, categorizing instances of `geopolitical:area` as `geopolitical:isSuccessorOf value X`, where *X* is another (geopolitical) area.

Novice ontologist assessment. There were two groups of students involved: Bc-level students in a course on Artificial Intelligence (AI) and MSc-level students in a specialized course on Ontology Engineering (OE). Both courses provided a certain degree of OWL modeling experience (in Protégé and Manchester syntax) prior to this exercise, although OE went into more depth as regards the underlying DL and reasoning. There were 17 AI students and 10 OE students altogether. In both courses the students were first provided with a 30-minute overview of the notions of CE (in $\mathcal{L}_{SE}$) and FC roughly as presented in Sections 2 and 3 of this paper. Then they completed an assignment consisting of 20 atomic tasks, all available in a single sheet of a web questionnaire.[32] In each atomic task the student was required to provide an answer to the question "Is the class CE a meaningful category for categorizing objects of class FC?", where FC was a named focus class and CE was a class expression in Manchester syntax. The answer was again collected on the 5-point Likert scale, with an additional option '*no judgment, since I don't understand the example*'.

---

[32]The questionnaire was in Czech. The English translation of a sample task is available in Appendix A.

The 59 tasks from the initial sample were randomly divided into three questionnaire versions (one **c4** task was used twice) to eliminate cribbing; the numbers of returned questionnaires per version were 7, 9 and 11, respectively, with a balanced proportion of AI vs. OE students. To avoid protracting and biasing the experiment, the students were instructed to only judge the CEs by the expression itself, i.e. without consulting the respective ontology specification or other external resources. However, specifically for the 'conference' domain of OntoFarm, they were provided with a brief domain glossary (since as students they were not expected to have experience with conference organization matters). In both sessions, 30' sufficed to all students for completing the (20-task) assignment.

We aggregated the results by questionnaire task, and then both by the course and by CE pattern. The aggregation was carried out by simple summation over the answer values rescaled to the $[-1; 1]$ interval (i.e. 'certainly' turned to 1, 'perhaps' to 0.5, both 'borderline' and 'no judgment' to 0, etc.), and then normalized by dividing by the number of students on the task. The normalized sum of the answer values of the *i*-th task is then

$$NS_i = \frac{\sum_{j=1}^{k} v_{i,j}}{k} \tag{12}$$

where $v_{i,j}$ is the rescaled answer value of respondent *j* for task *i*, and *k* is the total number of respondents of task *i*. For example, a task assigned to eight students, with the answers 'certainly', 'perhaps', 'borderline' and 'perhaps not', all present twice, yields

$$NS = (2*1 + 2*0.5 + 2*0 + 2*-0.5)/8 = 0.25$$

A short digest of the results follows:

– The average NS over all 60 tasks was 0.07, i.e. rather low, although positive. Of the 60 NS values, 28 were positive, 5 zero and 27 negative. The values strictly below 0.25 and above -0.25, possibly viewed as 'borderline aggregates', were 34 (57%).

– Perhaps most important, the average NS was highest for **c4** (0.21, with 15 positives, 1 zero and 7 negatives), lower for **c3** (0.02, with with 9 positives, 3 zero and 8 negatives), and lowest for **c2** (-0.05, with with 4 positives, 1 zero and 12 negatives).

– The cases[33] with highest positive and lowest negative values are in Table 6; the CE pattern is listed in the third column. We see that the cases with highest positive polarity tend to achieve higher absolute values than the cases with highest negative polarity, i.e., it was more likely for the students to agree on the positive than on the negative cases. We also see that **c4** dominates the upper end of the spectrum. Interestingly, the negative cases correspond each to a different CE pattern and also have different semantic roots: the 'village with baseball player as head-chef' one has already been discussed before (distant and dubious inference), the 'conference in city' one deals with a seemingly mandatory property leading to $Cat \equiv FC$ (here, however, the experts' consensual opinion diverged: how about future editions not yet having a location, or virtual conferences?), and the 'day followed by Friday' only holds for one individual, in turn.

– The average NS was higher for the OE students (0.12) than for the AI students (0.04), which might be attributed to more developed 'ontologistic thinking' of the former. The *inter-task* variance, indicating the tendency towards giving uneven values (averaged over the students filling the same task) across the questionnaire, was about the same (0.16) for both courses. However, the *intra-task* variance, indicating the degree of disagreement among the students filling the same task, was higher for the AI students (2.51) than for the OE students (2.12), i.e. the rating of the latter was more coherent.

In comparison with the 'expert ontologist' assessment:

– The students gave a significantly lower score: only about a half of the tasks had a positive NS, compared to 92% (54/59) in the final consensus of experts. This can be explained by their lower ability to figure out specific situations in which less obvious categories might become meaningful.

– If we apply the same method of average NS computation on the initial assessment of experts, the proportion of 'borderline aggregates' between -0.25 and 0.25 is only 14% (in contrast to 57% for the students' values).

– There is agreement on the less frequent 'meaningfulness' of **c2** CEs (i.e., on the lower reliability of the FCE pattern **p2**). Out of the 17 respective tasks, as mentioned above, only 4 were viewed as 'meaningful' by the students and 13 by the experts (who in turned judged all tasks with other patterns, except one, as 'meaningful').

– As regards the case-by-case comparison between the students and the experts, there is also a correlation in the sense that the 43 experts' clear positives obtained a positive average NS from students (0.14), while the 14 initially 'clash' cases obtained a slightly negative average NS (-0.07) and the 3 negative cases obtained a clearly negative average NS (-0.24).

7.2. Summer 2018 campaign

In the second campaing we tried to modify the setting so as to avoid some biases and gaps appearing in the first campaign, in particular:

– The even distribution of tasks between LOV and OntoFarm was judged inadequate, as OntoFarm is by an order of magnitude smaller, addresses one domain only, and its ontologies have been created artificially, even if based on real-world non-ontological resources.

– Assessing the CEs solely based on their formal representation risked of suffering from a comprehension bottleneck.

There was no a priori expert assessment this time (assuming that the correlation of the expert and novice assessment had been adequately studied in the first campaign).

Initial sampling and task preparation   This time, all tasks were generated from the LOV. In contrast to the 2016 campaing, we also added CEs matching the pattern **c1**, to verify the assumption that named subclasses are by default meaningful categories. We first pre-selected LOV ontologies that satisfied the following two conditions:

– they had more than 90% of their classes equipped with the `rdfs:label` value (since we would need these values for the questionnaire)

---

[33]Most namespace prefixes used can be expanded using the `prefix.cc` service. Prefixes unlisted by this service follow: **p-act**=http://purl.org/procurement/public-contracts-activities#, **p-aut**=http://purl.org/procurement/public-contracts-authority-kinds#, **pl**=http://www.loc.gov/premis/rdf/v1, **pl-sm**=http://id.loc.gov/vocabulary/preservation/storageMedium#, **sigkdd**=http://oaei.ontologymatching.org/2016/conference/data/sigkdd.owl

Table 6

CEs with highest and lowest average NS of student scores, 2016 campaign

FC	Subcategory	CE pattern	Avg. NS
ofrd:FridgeFreezer	ofrd:styleOfUnit value ofrd:SingleDoor	**c4**	0.91
gr:BusinessEntity	pco:mainActivity value p-act:GeneralServices	**c4**	0.86
gr:BusinessEntity	pco:authorityKind value p-aut:LocalAuthority	**c4**	0.61
akt:Generalized-Transfer	akt:information-transfer-medium-used value akt:Email-Medium	**c4**	0.59
p1:Storage	p1:hasStorageMedium value p1-sm:mag	**c4**	0.59
fabio:Item	fabio:isStoredOn value fabio:web	**c4**	0.50
...	...	...	...
dbo:Village	dbo:headChef some dbo:BaseballPlayer	**c3**	-0.50
sigkdd:Conference	sigkdd:City_of_conference some Thing	**c2**	-0.56
gr:DayOfWeek	gr:hasNext value gr:Friday	**c4**	-0.56

– they had at least 10 classes (to eliminate the long tail of very small ontologies).

The actual sampling was then performed on approx. 130 thousand CEs generated from the 72 ontologies that satisfied the above conditions. From this pool we randomly sampled ten tasks for every CE pattern (**c1**, **c2**, **c3**, **c4**). No manual filtering of tasks was applied this time, thus entities with cryptic IRIs were allowed (relying on the availability of the `rdfs:label` values). In addition to the labels, we also collected from the ontologies:

– *verbal definitions* of the involved entities, as the `rdfs:comment` values or even as definitions manually found in the documentation of the ontologies
– selected *axioms* in which the entities appeared in the RHS.

Unfortunately, the sampling results exhibited some potentially undesirable features, and we did not have time to redesign the sampling because of the planned experiment dates (within the schedule of both courses) that we were unable to shift. Namely:

– Some domain ontologies contained links to upper-level ontologies. If the FC was then picked from an upper-level ontology, it was highly abstract (e.g. 'Feature', 'Object' or 'Endeavor'), and its relationship to domain-specific concepts of the CE was hard to figure out. The assessment then had 'strong philosophical flavor', and the setting was unrealistic wrt. our target use case, since upper-level entities would not typically be sought as focus classes when publishing linked datasets.
– One of the tasks referred to an ontology in a language different from English (namely, Spanish).

By the results and the students' feedback it however does not seem that these infelicities would have seriously biased the experiment.

The structure of the CEs was *verbalized*, using simple NLP patterns plus occasional manual tweaking to assure grammatical correctness of the generated sentences. For example, if the property label P in a CE matching the pattern **c2** was a noun phrase, the CE was verbalized as "A specific kind of **FC**: such that has a **P**". On the other hand, if the property was labeled by a verb phrase, the verbalization of the CE became "A specific kind of **FC**: such that **P** something." The verbalization for the other compound CE patterns was analogous, e.g., for **c3**, ". . . such that has a **P** that is a **C**" (for a noun phrase P) and ". . . such that **P** a **C**" (for a verb phrase P), respectively.

Novice ontologist assessment. There were, again, two groups of students involved, from the same courses as in 2016 (Bc-level AI and MSc-level OE). The amount of prior training in OWL was also similar as in the 2016 campaign. There were 15 AI students and 16 OE students altogether. In both courses the students were first provided with a 30-minute overview of the notions of CE (in $\mathcal{L}_{SE}$) and FC roughly as presented in sections 2 and 3 of this paper. Then they completed an assignment consisting of 8 atomic tasks (two for each of the four CE patterns), all available in a sin-

gle sheet of a web questionnaire.[34] The time allocated to completing the questionnaire was 40 minutes, i.e. 5 minutes per task.

The task question was slightly modified: it explored to what the degree the category is meaningful and *reusable*. The rationale was that possibly even subcategories with very small absolute or relative frequency might be viewed as meaningful (this term being rather vague and subjective), but undoubtedly their reusability should be perceived as low. Besides, the novelty of the assessment task setting compared to the 2016 campaign was in the following:

- The questionnaire separated the meta-question on *comprehension* from the actual 'meaningfulness' assessment, for each task. A separate question now inquired to what degree the student understood the meaning of the *individual entities* (assuming that the lack of familiarity with the entities strongly impacts the competency to assess a compound CE), with possible values that can be shortened as: 'quite familiar', 'roughly', 'pretty vague idea' and 'no clue'.
- The students could textually *justify a negative value*.
- For the compound CEs, the students could provide a *noun phrase* to which the verbalization of the category could be compressed.

We will however not discuss the last two types of metadata elements (the unstructured ones) in the current paper, to avoid thematic dilution.

In addition to the FCP tasks assessment, the questionnaire also examined the students' assessment of their own level of *written English*.

We computed the normalized sum (NS) of the task assessment values, as in the 2016 campaign (using Equation 12). The core results are as follows:

- The average NS over all 40 tasks was 0.27, i.e. much higher than in the first campaign (0.07). Of the 40 NS values, 34 were positive, 2 zero, and only 4 negative. This can possibly be attributed to the longer time available for each task, to the higher amount of available documentation, and/or to the verbalization of the CEs.

- The relative position of the compound CE patterns did not change from the 2016 campaign. The average NS was 0.33 for **c4** (all ten NS values being positive), 0.30 for **c3** (8 positives, 1 zero and 1 negative) and 0.16 for **c2** (7 positives, 1 zero and 2 negatives). The NS for **c1** was the highest, 0.39. This value is still surprisingly low, it is however strongly influenced by a single task with a negative NS. Upon removing this outlier, the NS would be 0.47.
- The cases with highest positive and lowest negative values are in Table 7; both the FC and the subcategory are now shown at the level of labels, just as presented to the students. The property and its filler, whether a class or an individual, are separated with a colon. The underlying ontology is referenced in the third column, through its nickname, which can be resolved against the LOV portal by appending it to `https://lov.linkeddata.es/dataset/lov/vocabs/`; to give a rough idea on the ontology size, the number of its classes is shown in parentheses. The CE pattern is listed in the fourth column, and the average NS value in the fifth column. Interestingly, the students were positive towards some highly abstract focus classes (Quantity, and even Thing). As regards the negative cases, the first three rather seem to do with too-far-reaching connections between the involved entities; the last case suggests some odd modeling in the ontology itself (the students commented that a subcontinent should rather contain a country than be contained in it).

We also computed the relative frequencies reflecting the impact of the (declared) *English writing skills* and of the *comprehension of entities* on the assessment value:

- Of the 168 assessments by students with excellent or very good English skills, 102 (61%) were positive ('certainly' or 'perhaps'); in contrast, of the 80 assessments by students with fair or basic English skills, only 37 (46%) were positive.
- Of the 176 assessments where the students comprehended the meaning of the CE entities ('quite familiar' or 'roughly'), 126 (72%) were positive ('certainly' or 'perhaps'). In contrast, only 13 assessments of 72 (18%) where the students did not comprehend the meaning of the CE entities ('pretty vague idea' or 'no clue') were positive.

---

[34]The questionnaire was again in Czech, except the CEs (verbalized in English, to avoid issues with the inflection grammar of Czech). The English translation of a sample task is available in Appendix B.

Table 7
CEs with highest and lowest average NS of student scores, 2018 campaign

FC	Subcategory	Onto (# cl.)	CE pattern	Avg. NS
Security mechanism	BlockCipher	stac (211)	**c1**	0.79
Object	Machine	pext (488)	**c1**	0.75
Sensor	Gyro sensor	ha (88)	**c1**	0.75
Social position	has title: Prime minister	pext (488)	**c4**	0.75
Quantity	Mass	schema (625)	**c1**	0.67
Document	cites: Protocol	vivo (146)	**c3**	0.67
Thing	has dimension: Thing	ecrm (84)	**c2**	0.64
...	...	...	...	...
Service	has forecast: Thing	km4c (582)	**c2**	-0.21
CRM entity	is separated from: Beginning of existence	ecrm (84)	**c3**	-0.29
Feature	Camper service	km4c (582)	**c1**	-0.33
Subcontinent	in country: Thing	swpo (57)	**c2**	-0.50

In order to reflect the degree of entity semantics comprehension in the assessment (with the assumption that more weight should be given to 'more informed' assessment), we also applied simple numerical weighting: the formula from Equation 12 was changed to

$$NS_i = \frac{\sum_{j=1}^{k} c_{i,j} \cdot v_{i,j}}{k} \qquad (13)$$

where $c_{i,j}$ was set to 1 if the comprehension of the entities of task i by respondent j was 'quite familiar', to 0.66 for 'roughly', to 0.33 for 'pretty vague idea', and to 0 for 'no clue'. The results did not dramatically change. The average NS per CE pattern became as follows: 0.35 for **c1**, 0.32 for **c4**, 0.25 for **c3** and 0.14 for **c2**. Compared to the non-weighted normalized sum, **c4** marked the smallest decrease of all CE patterns.

7.3. Summary

The role of the cognitive experiments was to eventually attempt to assess the reusability degree of different individual CEs and thus (indirectly) their patterns, which is tied to the central idea of the whole approach (in which, for example, the FCE patterns are merely instrumental).

Across the different campaigns and settings, the order of the CE patterns according to the average 'meaningfulness' of the categories remains stable: (**c1** >) **c4** > **c3** > **c2**. We can also conjecture that the categories are more likely to be viewed as 'meaningful' if the assessor is an expert, or if some of the following holds: more information is available on the entities and context; the categories are better presented (e.g., verbalized); more time is available per category (task).

Data availability The data from both experiments is available on the web through `https://bit.ly/FCP_cogex`. The page contains, for each of the campaigns, the links to:

– The description of each task (consisting in evaluation of the meaningfulness of a class expression) in each questionnaire variant
– The table with calculation of aggregated results.

8. Discussion

Since the empirical part of the paper may appear a bit fragmented and the results hard to align, in this section we first provide an integrative meta-view of the surveys / experiments settings and results.

From this we depart to a discussion of limitations and open questions of the analysis.

8.1. Meta-view of the empirical analyses

In Table 8 we synoptically summarize the three empirical pillars of our research so far, as elaborated in Sections 5, 6 and 7. We see that the surveys/experiments are to a large degree complementary, differing in their features: in the *structural type of the data source*; in the focus on either *direct* analysis of CEs or on their underlying patterns that are only *indirectly* tied to the CEs; in the *objectivity/subjectivity* of the obtained data; and, finally, in the actual *scope of the CE and FCE patterns*, and of the *focus classes* examined. Notably, all data sources currently refer to the Tbox. As mentioned before (esp. in Section 2.1), the fourth pillar of the empirical analysis of the FCP problem should be

the analysis of the CEs occurrence in the *RDF datasets Abox* (which is ongoing but did not fit into the current paper).

The last row in the table attempts to summarize the core findings of each analysis. At the first sight the arrangement of the CE/FCE patterns might look incoherent. Especially, **c2** is (by Sec. 5) very rarely used in the axioms RHS, which contrasts with the fact that we can generate a huge number of matching CEs (by Sec. 6) through FCE pattern **p2** and the proportion of these CEs that are 'meaningful' is not dramatically lower than for the other CE patterns (by Sec. 7). This paradox can however be explained as follows:

- Sec. 5 considers the CEs within the RHS of axioms, where they primarily serve as a means for inferring the subordination of *arbitrary* instances (or classes) to the named class appearing in the LHS of the axiom; it is thus desirable that the RHS would specify some restrictive filler (as in **c3**) and not just `owl:Thing` (as in **c2**), to avoid 'false positives'.
- On the other hand, Sections 6 and 7 already study the categorization in the setting with a known *focus class*, i.e. not for arbitrary instances. Then even the categorization of individuals based solely on the property they appear together with, irrespective the filler (as in **c2**), might work.

As regards the high ranking of **c4** in the cognitive experiments, in contrast with its low ranking in the other two studies, there is no discrepancy at all, since the cognitive experiments analyzed the meaningfulness of the categories (and thus, indirectly, their CE patterns) and not the frequency of occurrence of CE patterns at all (they were sampled evenly). The findings are thus completely orthogonal: there will typically be few **c4** CEs at hand, but they should be counted with a high weight in the FCP computation. Recalling the results from Section 6.2, for focus classes lower in the taxonomy, even the quantity of **c3** and **c4** CEs may proportionally increase.

8.2. Limitations and open questions

Since the paper explores a substantially novel problem space, the coverage of its different corners is still rather limited. In this section, we discuss three limitations and/or open questions, in turn: the omission of Abox data in the whole process, the simplification made in the empirical analysis of CE pattern frequency, and the reliance on particular Tbox design principles in the application of FCE patterns.

The presented research currently ignores the role of *Abox analysis*, both in the *actual FCP computation* (within a concrete reuse problem) for a specific ontology as well as in the *inductive process* of CE pattern weight estimation. The former means that the presented approach is tuned towards the 'cold-start' setting (new ontologies not yet referenced in data); otherwise the relevance of concrete CEs could be estimated (though with some computational cost) from existing datasets instantiating them. The omission of the latter (which would consist in generating CEs for a given CEL over a representative set of 'training' ontologies, measuring their frequency of instantiation in data, and aggregating those frequencies by CE pattern) to date does not have any principled reason, and is only owing to limited human resources on the project.[35]

The computation of the *CE pattern frequency* in axioms from Sec. 5 targeted the top-level structure in the axiom RHS (with additional distinction of the CE patterns from $\mathcal{L}_{SE}$). How complex are the embedded expressions within this top structure remains undistinguished. For example, for the conjunction, it would make sense to distinguish between a conjunction of two named classes, a conjunction of two property restrictions, a conjunction of a named class and a restriction, and several others. Characterizing the utility of a CE pattern based on the top-level structure alone is obviously superficial, as it would vary according to the embedded levels.

The design of our *FCE patterns*, where *global* restrictions (domain+range) are used in the pattern template and *local* (existential) restrictions in the validity constraints, puts a strong assumption on the engineering approach used in the ontology analyzed. If the ontology lacks global restrictions[36] then the FCP will be under-estimated, since the FCE patterns will not match it at all. On the other hand, if the ontology lacks local restrictions, the FCE patterns will possibly generate many useless categories (those equivalent to the FC, or with empty or singleton extension), which would then have to be manually filtered out. This caveat could be compensated by computing the overall proportions of the number of classes plus local and global restrictions

---

[35]While some early experiments are under way, their inclusion (in a low-maturity state) would have made the current paper lengthy and the contribution diluted.

[36]Note that some best practices, e.g., in biomedicine, used to argue against domain/range axioms, see e.g. [18]

Table 8

Summary of the complementary surveys / experiments, with core findings (last line)

Survey / experiment	CE pattern frequency in axioms (Sec. 5)	FCE pattern occurrence in the Tbox (Sec. 6)	CE assessment via the cognitive experiments (Sec. 7)
Data source type	RHS of axioms (OWL)	Axioms (primarily RDFS)	Axioms (primarily RDFS); human assessors
Source ontologies	LOV, BioPortal, OntoFarm	LOV, OntoFarm	LOV, OntoFarm (2016 campaign); LOV (2018 campaign)
CEs examined	Directly	Indirectly (through FCE patterns)	Directly (by humans), but generated using FCE patterns
Objectiveness	Objective	Objective	Subjective (human judgment)
CE/FCE patterns	**c2, c3, c4**	**p1, p2, p3, p4, p5**	**c1, c2, c3, c4**
Focus classes	(Not relevant)	All classes	59+40 randomly selected classes
Gist of the quantitative findings	CE frequency: **c3** most frequent (by an order of magnitude)	FCE pattern occurrence: **p2 > p1 > p3 > p4 > p5** with **p1** and **p3** being swapped for less general FCs	Proportion of CEs meaningful for human assessors: **c1 > c4 > c3 > c2** (but not differing dramatically)

in the examined ontology and then adjusting the FCP accordingly.

Another potentially arguable assumption under which we gave preference to the *existential property restriction* as the cornerstone of our initially chosen CEL, over the *Boolean conjunction* (which would have been another simple candidate for a CE pattern), is that we *cannot* rely on proper equipment of ontologies with *disjointness axioms*. As mentioned in Section 5, in cases where we can (through pruning the ontology using those axioms) single out truly compatible pairs of classes, the consideration of conjunctive categories would be promising; in other cases it would however lead to explosion of output CEs, which we wanted to avoid.

9. Operationalization and prototype implementation

While most of this paper is devoted to the description and formalization of the FCP 'view' of ontologies and to the associated empirical analyses, we also provide a tentative (or rather, illustrative) operationalization of the empirical findings into FCP computation weights, an describe an early prototype of an ontology search (for reuse) tool leveraging on FCP.

9.1. Tentative operationalization of the empirical results

We understand the obtained insights into the usage of OWL class expressions and their perception by humans in general as a research achievement per se.

However, the starting point for the overarching empirical study was an 'engineering' goal (possibly modest compared to the extent of the performed surveys and experiments): to propose *adequate weights* for the CE patterns to be used when computing the FCP in the context of an ontology reuse scenario.

In these terms, based on the cognitive experiments in particular, we can see that **c4** (individual value restriction) and to some degree **c3** (existential restriction) may successfully complement **c1** (named class) when harvesting categories from a reused ontology. The survey of the CEs in axioms, in turn, suggests that the chosen CE language ($\mathcal{L}_{SE}$) covers a decent part of the space of relevant languages and could be applied in the first, albeit rough approximation of an ideal FCP formula. Thus, if a requirement came to design an ontology reuse metric leveraging on the FCP, based on the cognitive experiments we could tentatively set the weights in Equation 4 according to the NS of the assessments in the cognitive experiments. Two simple options of generating weights from the assessment NS are quantified in Table 9: by transforming the average NS from the $[-1; 1]$ interval to the $[0; 1]$ interval, and by computing the ratio of positive NS values of all NS values for the given CE pattern.

Considering that the lowered average NS values of **c1** in the 2018 campaign were presumably caused by randomly sampled highly general classes, which would not normally appear in the focus class role (as in the `Feature` class in Table 7), we should probably set $w^P(\mathbf{c1})$ to 1. The remaining weights can be set, heuristically reflecting both calculation methods and the results of both campaigns in combination, for example as $w^P(\mathbf{c2}) = 0.5$, $w^P(\mathbf{c3}) = 0.7$, and $w^P(\mathbf{c4}) = 0.8$.

Table 9

Alternative CE pattern weights derived from the cognitive experiments

	c1	c2	c3	c4
Average NS mapped to [0, 1]				
Campaign 2016		0.48	0.51	0.61
Campaign 2018	0.70	0.58	0.65	0.67
Campaign 2018 weighted	0.68	0.57	0.63	0.66
Ratio of positive NS values				
Campaign 2016		0.24	0.45	0.65
Campaign 2018	0.90	0.70	0.80	1.00
Campaign 2018 weighted	0.90	0.80	0.90	1.00

This would yield the following FCP estimate formula, to be applied on an ontology O and focus class FC, the language $\mathcal{L}$ being $\mathcal{L}_{SE}$:

$$\widehat{FCP}(O, FC, \mathcal{L}, \{\{\mathbf{p1}\}, \{\mathbf{p2}\}, \{\mathbf{p3}\}, \{\mathbf{p4}\}\}) =$$
$$Occ(\{\mathbf{p1}\}, FC, O) \cdot 1 + Occ(\{\mathbf{p2}\}, FC, O) \cdot 0.5 +$$
$$+ Occ(\{\mathbf{p3}\}, FC, O) \cdot 0.7 + Occ(\{\mathbf{p4}\}, FC, O) \cdot 0.8$$

The current version of the OReCaP tool (see the next subsection), when launched, proposes these values, i.e. $w^P(\mathbf{c2}) = 0.5$, $w^P(\mathbf{c3}) = 0.7$, and $w^P(\mathbf{c4}) = 0.8$. The formula should however only be taken as a very first, tentative initial setting (to be adapted by the user as needed) and not as a granted recipe.

9.2. OReCaP prototype

To demonstrate the whole focused categorization framework, specifically for $\mathcal{L}_{SE}$, patterns **p1**–**p4** and in the context of the ontology reuse scenario, as well to make it possible for both us and any interested subjects to perform experiments on various input ontologies and focus classes, we implemented a web-based application named OReCaP.[37] The application is available online at `https://fcp.vse.cz/orecap`. A demo of its earlier, limited version was already presented at the ISWC 2020 conference [16]. What is novel in the currently available version compared to that in the demo paper is, in particular, the support of *actual ontology reuse*, mediated by the user's interaction with different categories constructed using different FCE patterns.

The interaction workflow with the tool consists of several, possibly iterative, phases:

---

[37]The acronym refers to the terms 'ontology reuse' and 'categorization power'.

1. The process starts with a *keyword-based search* where the input consists of at least one *focused class keyword* and of optional *additional keywords*. The intuition behind the search interface is that the focused class keyword/s denotes the high- or medium-level type/s of entities whose instances are to be further sub-categorized using concepts from the ontology; the additional keywords, on the other hand, correspond to whatever domain terms. Imagine, for example, that the data is currently stored in a relational database. The focused class keyword might then often be the name of the top-level table (which can be, e.g., 'Client', 'Patient', 'Vehicle', 'Account', or the like); the additional keywords can be taken, e.g., from the names of subordinate tables, table columns, or predefined values for the fields.

2. The search returns a *sorted list of ontologies* whose classes match one or more of the provided keywords by their IRI, name or description; classes with a match of focused class keyword are listed first. The matched classes are listed for each ontology. Classes that match the *focused class keywords* are preselected (i.e., checked) by default; classes that matches the *additional keywords* are not preselected but can be selected (checked) manually by the user.

3. The next step is to execute the *FCP calculation* for a chosen ontology, given the selected classes as focus classes, by clicking on the 'Calculate FCP' button. In a pop-up window, metadata about the ontology including its IRI and namespace is displayed, along with the total FCP score, which is calculated based on the FCP weight values and the categorizations (i.e., class expressions) listed at the bottom. This score is the sum of all partial scores for each focus class. The weight values can be adjusted for each calculated ontology according to the user's assessment of each CE pattern, and the resulting FCP score will change accordingly. The global FCP weights can be changed in the settings section, so that every new FCP calculation would use them as the default weight values.

4. The calculated FCP score is then displayed below the ontology overview in the search results, and also saved to a *comparison list*, which shows the FCP-based ranking of the ontologies.

5. From the comparison list it is possible to proceed to the *reuse summary* view of a given ontology. There the user can inspect all generated

categories in the form of a pop-up tree, and can check those s/he eventually wants to reuse.

6. The actual *reuse* is currently performed two ways: first, a particular ontology (chosen by the ranking) can be downloaded as is. Second, the axioms defining the named classes newly constructed from compound expressions can be downloaded separately from the reuse summary view, to be later manually added to the ontology.

Let us assume the user wanting to publish data about business contracts and their payment, and seeking a suitable ontology for their subcategorization. A partial screenshot showing the overall search results together with FCP scores for two ontologies is in Fig. 2. The focus class keywords (here, just one) are entered in the top-left field; their matches are therefore proposed (hightlighted in blue) as focus classes, by the tool. The user has also provided additional keywords, which may improve the result ranking but do not produce further focus classes (unless the user pro-actively highlights them, too).

For the first ontology, PPROC, a snippet of the reuse summary window is then shown in Fig. 3; the numbers (e.g., "(1/5)") indicate how many categories were chosen in the given sub-tree. The user has chosen two categories, corresponding to 'contract that has been anyhow modified' (CE pattern **c2**) and 'contract that has been ended by procedure resignation' (CE pattern **c3**). The set of axioms defining new named classes based on such compound CEs, with shortnames currently created by trivial concatenation, can then be downloaded; for example:

$$\texttt{ContractHavingContractModification} \equiv$$

$$\texttt{Contract} \sqcap \exists \, \texttt{ContractModification} \, . \top$$

A custom namespace can be set for those new classes.

OReCaP makes use of the Linked Open Vocabulary API[38] for the keyword-based search and for retrieving the ontology metadata. The FCP calculation itself is implemented on top of OWL API, in combination with the jFact reasoner.[39] OWL API is used to load and parse the ontology source codes, and jFact is used to infer class expressions. Both the front-end and back-end implementation of OReCaP are open-source and available on GitHub[40] under the MIT license.

10. Related work

Since the research described in this paper addresses the focused categorization power problem from various angles, multiple areas of related research can be identified. In this section we report on the following, in turn: abstract notions similar to our notion of FCP; empirical studies on presence of class expressions and structural patterns in ontology repositories; cognitive experiments on assessing ontological structures; concept learning in DL; ontology reuse metrics and methods.

10.1. Abstract notion of focusing or categorization (power) in ontologies

We are unaware of prior work on the same topic of FCP as we coin it in the current paper. We will however reference some related research that overlaps with ours at the abstract level.

The notion of focusing recently appeared in the work of Gogacz et al. [9]. The so-called focusing solution pairs a set of predicate symbols that describe a database schema (that is, a set of predicates) with a set of assumptions on the partial completeness on the data and the ontology (closed and fixed queries). In their approach, focusing is about choosing which parts of the data and ontology are to be declared complete, to allow for efficient reasoning. In our approach, focusing is about choosing ontology classes for whose instances in data we would like to obtain many meaningful categorization options; the categorization itself however need not rely on logical reasoning (never mind using the ontology), but can be based on whatever kind of classification model or even made by humans.

The term classification/categorization power previously appeared in many scientific texts, however, rarely as a rigorously defined notion. For example, on many occasions, automated classifiers (typically, machine-learning-based) are reported to have certain 'classification power' with respect to classes from an ontology, which is merely an informal circumscription of measures such as accuracy or error rate. The 'power' also clearly pertains to the *classifier* and not to the on-

---
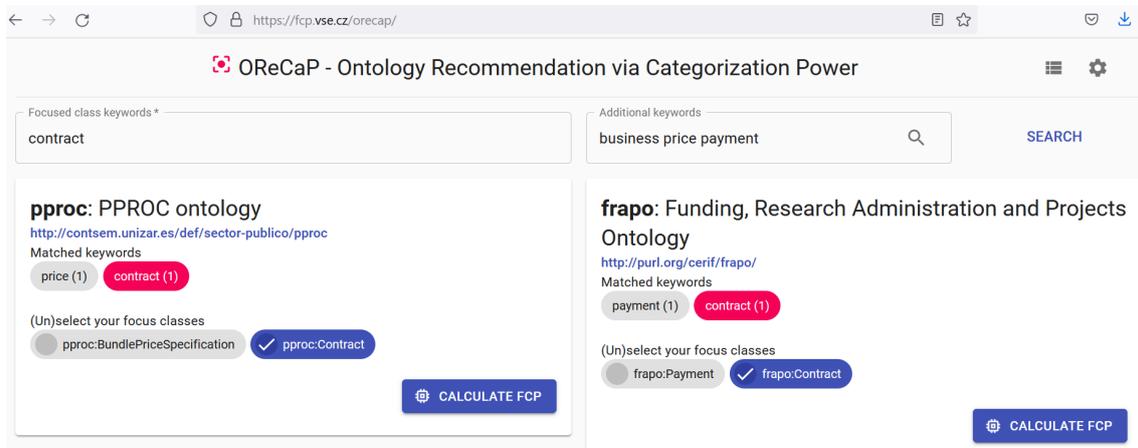
Fig. 2. OReCaP interface: two ontologies found via keyword search, with focus class and additional match

Fig. 3. OReCaP interface: the reuse summary, with user's choices of two CEs

tology. The association of the notions of 'categorization' is thus merely verbal.

Partially relevant is the analysis made by Giunchiglia & Zaihrayeu [8], who categorized 'lightweight' ontologies with respect to two dimensions: complexity of labels (simple noun phrases vs. use of connectives and prepositions) and use of 'intersection' operator allowing to combine atomic entities of different nature (e.g., the atomic concepts 'Italy' and 'vacation' implicitly combine into 'vacation in Italy'). Maximal 'classification power' is obtained when both explicitly complex labels and implicit concept combinations are allowed.

This however only applies to classifying documents extrinsic to the ontology, since 'intersection' of concepts of different nature is not coherent with the set-theoretic semantics of DL. Overall, their 'classification power' is a global property of the method by which the ontology has been built. In contrast, our notion of FCP applies to individuals intrinsic to the DL world of the ontology and is calculated with respect to a focus class.

A. Rector's work on entangling hierarchies (normalization) [17] addressed a different problem than us, but to some degree analogously considered the compound

concepts as an alternative to named ones. This applies to 'partitioning' or 'refining' concepts, that only modify the 'self-standing' concepts; secondary partitioning aspects should not be expressed through subclassing (yielding a multi-hierarchy) but through existential restrictions filled with classes from separate 'codelist' taxonomies. For example, a class `SteroidHormone` would not have two parents, `Steroid` and `Hormone`, but only the first one, while the latter subordination would be made through an existential restriction: `SteroidHormone ≡ Steroid ⊓ ∃ playsRole.Hormone`.

Our own ongoing work on the PURO modeling language [24] deals with various options how the same 'background' state of affairs can be expressed in OWL. PURO structurally resembles OWL but relaxes some of its modeling constraints. A library of transformation patterns allows to proceed from one PURO model to alternative OWL ontologies in different encoding styles. An example relevant to our case is the notion of *enterpreneur*, which is likely to be expressed as type in PURO, but could be translated to relationship (`insuranceCategory`) restricted to the `Entrepreneur` individual in OWL (i.e. a compound CE), assuming we prefer an encoding style using object properties with 'codelist individuals'. Analogously, *born in* may possibly be a relation in PURO but can be translated not only to OWL property restrictions but also to named classes such as `PersonBornInUK`, assuming we prefer an 'encapsulating' encoding style used, e.g., in the DBpedia Ontology.[41] Modeling in PURO and applying the transformation patterns may thus make hidden (compound) relevant concepts explicit in the domain as atomic concepts. A similar account of alternative 'typecasting' beyond named classes (but with smaller coverage) has been given by Krisnadhi [11].

Certain research in cognitive psychology might also be relevant wrt. the notion of FCP. In particular, the notion of *graded structure* of categories [6] can be applied to our concept pattern weighting. The authors suggest that there is "a basic level of abstraction (e.g. CHAIR, DOG), . . ." which is "further discriminated at the subordinate level (e.g. KITCHEN CHAIR, SPANIEL) and abstracted at the superordinate level (FURNITURE, ANIMAL)" [6]. Presumably, this basic level might often coincide with the focus classes to be subcategorized in practical settings. Furthermore,

---

[41] `http://wiki.dbpedia.org/services-resources/ontology`

the graded structure may include "ad hoc, goal derived categories such as GOOD PLACES TO HIDE FROM THE MAFIA" [6], analogous to our CEs defined via an existential restriction.

10.2. Empirical studies of ontology features

As regards the analysis of ontology repositories in terms of various aggregated features and metrics (logical, graph, lexical etc.), there has recently been renewed interest, following up with the early work of Tempich et al. [25] (aiming to build a benchmark for testing ontology tools). A large scale study of OWL ontology metrics was carried out by Matentzoglu et al. [15]. However, the categorization power of ontologies has not been, to our knowledge, studied, never mind with the flavor presented here.

Our study on class expression frequency in axioms from Section 5 looks similar to the recent study carried out in the MontoloStats project [14]. Both studies essentially analyze the same ontology repositories (primarily, LOV and Bioportal), and refer to the suitability of ontologies for reuse. There is however a difference in the restrictions coverage. For an unclear reason, the MontoloStats study does not cover existential restrictions (which are central for our study, and also shown as empirically very frequent) at all, nor the conjunctive concepts. On the other hand, it covers (subclass axioms with) named class in the RHS, and also property axioms such as domain/range or functional property. Notably, even over the common subset of CE patterns in restrictions (such as disjoint, universal and cardinality restriction CEs) that some differences in the computed ranking appear between MontoloStats and our research; these may be due to additional distinct features in the methodology used.

Yet another stream of empirical research aims to study ontologies not on their own but from the point of view of LOD datasets in which they are used. This was the subject of the project by Asprino et al. [3], which produced a condensed representation of the global, virtual, 'LOD ontology' in the form of so called equivalent set graphs. Various metrics related to the connectedness and extensional size of ontology entities were computed; while this research does only addresses compound concepts, it is in line with our ongoing activity in analyzing the Abox imprint of (named as well as compound) class expressions.

10.3. Cognitive experiments on assessing ontological structures

Several cognitive studies using ontologies as material have been published in the ontological engineering research. However, they primarily address the capability of humans themselves to carry out the categorization of objects to a set of classes or to understand the structure of OWL expressions. A recent example of the former is a study on classifying domain entities to upper-level ontology classes [22]. An example of the latter is an earlier study on the human capability of deriving useful information from differently verbalized OWL statements [26]. Our research in Section 7 of this paper differs in that the humans were to assess the automatically build concepts as more or less plausible, thus generating ground truth. (Semi-automated) verbalization was present, too, but only played an auxiliary role, the actual subject of assessment being the formal CEs themselves.

10.4. Concept learning in DL

The heuristic construction of compound CEs from the ontology axioms, triggered by the identification of a focus class, bears some resemblance to *concept learning* in DL [13], which is triggered by the specification of the target concept. However, (supervised) concept learning aims to identify CEs that best approximate a concrete goal concept as whole. In contrast, CEs logically equivalent to the focus class are uninteresting as its subcategories, while categories that only cover a few per cent of the FC instances may be valid means for partitioning its instances. Moreover, concept learning fully depends on presence of instance data for the *particular concept* to be learned. Yet, for a newly designed ontology such data might not be available (the well-known cold start problem). Finally, the overall goal is different: while concept learning primarily aims at enriching the ontology Tbox with new axioms [12], in our approach the prime task is to reuse the ontology (or its fragments) 'as it is', and the formation of new equivalence axioms is an optional, secondary step carried out with human interaction.

10.5. Ontology reuse metrics and methods

The broad context of our research, the task of ontology reuse, was studied by Schaible et al. [20]: the users expressed their preferences on reuse strategy in a survey. The results indicate that reusing multiple entities from the same vocabulary may often be preferred; this corroborates the relevance of our approach to measuring the categorization power of ontologies with respect to focus classes.

Vocabulary reuse techniques similar to the use of FCP-based metrics also appeared in a recent project on combining popularity metrics with the credibility of the vocabulary designers [21]. As regards the designer credibility, this is a feature of the ontology itself similarly to FCP, but it is completely orthogonal.

Reuse support [7] is also systematically sought by the maintainers of LOV [27], primarily at keyword relevance level; we are in contact with them and will seek to integrate our complementary approaches.

11. Conclusions and future work

Ontologies are an important means of subcategorizing entities already known to belong to a general focus class. Ontologies with the best *categorization power*, in terms of the number and quality of available subcategories of the *focus class/es*, should have preference in *ontology reuse* scenarios. We demonstrated that the scope of subcategories need not be confined to named classes but can also cover *compound ones*. In the first approximation, we treated the computation of the focused categorization power of ontologies, beyond named subclasses, in terms of *simple existential restrictions over properties*. This appears particularly relevant for publishing datasets on the linked data cloud, which is relatively 'property-centric'.

Ongoing research addresses the analysis of the CE *instantiation in linked datasets* (Abox). This information can be employed two-ways. First, directly for the CEs generated from the Tbox of the given ontology considered for reuse. Second, the occurrence of instantiation of a compound CEs can be compared to the occurrence of its constituent entities, in the same dataset. The ratio of the compound vs. individual instantiation can then be aggregated per the pattern of compound CE, and provide an empirical grounding for the 'meaningfulness' of the CE pattern, complementary to the human feedback from the cognitive experiments. A dedicated thread of the Abox analysis will also study the availability of external, deferefencable *SKOS codelists*, which would enter the FCP computation through FCE patterns such as **p5**.

Another area of ongoing research concerns the techniques of *pattern-based lexical transformation* of compound CEs to named ones (now merely as a trivial con-

catenation with `having` in OReCaP). Based on an already performed empirical analysis of naming patterns in the LOV ontologies, the transformation method will be finalized and implemented; a naming pattern catalog will also appear, as a side-product.

In middle term, we plan to extend the *concept expression language* considered, as regards the theoretical analysis, empirical studies, and support by a next version of the OReCaP prototype. A simple extension to $\mathcal{L}_{SE}$ could be the inclusion of the *inverseOf* predicate: in some modeling styles only one of the pair of mutually inverse properties is present in the ontology, and the entity categorization should then be carried out even against the direction of the respective property. Next, we plan to move towards some other *concept expression patterns* identified as common in OWL axioms, by our study, such as universal restrictions, conjunctions (for ontologies equipped with disjointness axioms), disjunctions (relevant for group categorization) and some nested expressions, especially concepts based on *property paths* of length 2. Besides we also want to consider *data property restrictions*; they are rare in ontology axioms, but possibly highly relevant for generating subcategories, since data properties abound, in particular, in markup-oriented vocabularies such as *schema.org*. Relevance for the FCP is quite obvious for Boolean-valued properties, but even numerical properties, where the categories would correspond to value intervals, should be taken into account.

Within the scope of the different CE patterns, the syntactic *FCE patterns* providing the CEs will also be extended (e.g., by considering the *subPropertyOf* relationship) and new heuristic alternatives possibly added. The algorithm for *pattern-based generation* of CEs will also be enhanced, primarily by applying post-pruning according to the pattern applicability constraints. Overall, the transition to more expressive CELs and thus more expressive patterns will of course mandate deeper *computational complexity analyses*.

The pool of testing subjects in *cognitive experiments* should be extended to external ontology engineering experts (on the top of students and internal experts, engaged in the previous experiments). The new experiments should also feature more carefully pre-selected (rather than random) assessment tasks, so at to provide targeted feedback to so far unclear constellations of focus class and subcategory.

Since the main foreseen practical application of the whole approach is the improvement of *ontology reuse*, we also plan to explore this task in full, considering the FCP-based approach merely as a single element to be combined with other state-of-the-art supporting techniques for ontology reuse.

Acknowledgments

References

[1] OWL 2 web ontology language profiles (second edition). W3C Recommendation, December 2012. URL `https://www.w3.org/TR/owl2-profiles/`.

[2] OWL 2 web ontology language structural specification and functional-style syntax (second edition). W3C Recommendation, December 2012. URL `https://www.w3.org/TR/owl2-syntax/`.

[3] Luigi Asprino, Wouter Beek, Paolo Ciancarini, Frank van Harmelen, and Valentina Presutti. Observing LOD using equivalent set graphs: It is mostly flat and sparsely linked. In Chiara Ghidini, Olaf Hartig, Maria Maleshkova, Vojtech Svátek, Isabel F. Cruz, Aidan Hogan, Jie Song, Maxime Lefrançois, and Fabien Gandon, editors, *The Semantic Web - ISWC 2019 - 18th International Semantic Web Conference, Auckland, New Zealand, October 26-30, 2019, Proceedings, Part I*, volume 11778 of *Lecture Notes in Computer Science*, pages 57–74. Springer, 2019. . URL `https://doi.org/10.1007/978-3-030-30793-6\_4`.

[4] Franz Baader and Paliath Narendran. Unification of concept terms in description logics. *J. Symbolic Computation*, 21:277–305, 2001.

[5] Mariano Fernández-López, María Poveda-Villalón, Mari Carmen Suárez-Figueroa, and Asunción Gómez-Pérez. Why are ontologies not reused across the same domain? *J. Web Semant.*, 57, 2019.

[6] Liane Gabora, Eleanor Rosch, and Diederik Aerts. Toward an ecological theory of concepts. *Ecol. Psychology*, 20, 2008.

[7] Nuria García-Santa, Ghislain Auguste Atemezing, and Boris Villazón-Terrazas. The ProtégéLOV plugin: Ontology access and reuse for everyone. In *The Semantic Web: ESWC 2015 Satellite Events*, pages 41–45, 2015. .

[8] Fausto Giunchiglia and Ilya Zaihrayeu. Lightweight ontologies. In Ling Liu and M. Tamer Özsu, editors, *Encyclopedia of Database Systems*, pages 1613–1619. Springer US, 2009. ISBN 978-0-387-39940-9.

[9] Tomasz Gogacz, Víctor Gutiérrez-Basulto, Yazmín Ibáñez-García, Filip Murlak, Magdalena Ortiz, and Mantas Simkus. Ontology focusing: Knowledge-enriched databases on demand. In Giuseppe De Giacomo, Alejandro Catalá, Bistra Dilkina, Michela Milano, Senén Barro, Alberto Bugarín, and

Jérôme Lang, editors, *ECAI 2020 - 24th European Conference on Artificial Intelligence, 29 August-8 September 2020, Santiago de Compostela, Spain, August 29 - September 8, 2020 - Including 10th Conference on Prestigious Applications of Artificial Intelligence (PAIS 2020)*, volume 325 of *Frontiers in Artificial Intelligence and Applications*, pages 745–752. IOS Press, 2020. .

[10] Ullrich Hustadt, Renate A Schmidt, and Lilia Georgieva. A survey of decidable first-order fragments and description logics. *Journal of Relational Methods in Computer Science*, 1 (251-276):3, 2004.

[11] Adila Alfa Krisnadhi, Pascal Hitzler, and Krzysztof Janowicz. On the capabilities and limitations of OWL regarding typecasting and ontology design pattern views. In Valentina A. M. Tamma, Mauro Dragoni, Rafael Gonçalves, and Agnieszka Lawrynowicz, editors, *OWLED*, volume 9557 of *Lecture Notes in Computer Science*, pages 105–116. Springer, 2015. ISBN 978-3-319-33244-4.

[12] Jens Lehmann and Lorenz Bühmann. ORE - a tool for repairing and enriching knowledge bases. In *Proceedings of the 9th International Semantic Web Conference (ISWC2010)*, volume 6497 of *Lecture Notes in Computer Science*, pages 177–193, Berlin/Heidelberg, 2010. Springer. .

[13] Jens Lehmann and Pascal Hitzler. Concept learning in description logics using refinement operators. *Machine Learning*, 78 (1-2):203–250, 2010.

[14] Sven Lieber, Ben De Meester, Anastasia Dimou, and Ruben Verborgh. MontoloStats - ontology modeling statistics. In Mayank Kejriwal, Pedro A. Szekely, and Raphaël Troncy, editors, *K-CAP*, pages 69–76. ACM, 2019.

[15] Nicolas Matentzoglu, Samantha Bail, and Bijan Parsia. A snapshot of the OWL web. In Harith Alani, Lalana Kagal, Achille Fokoue, Paul T. Groth, Chris Biemann, Josiane Xavier Parreira, Lora Aroyo, Natasha F. Noy, Chris Welty, and Krzysztof Janowicz, editors, *International Semantic Web Conference (1)*, volume 8218 of *Lecture Notes in Computer Science*, pages 331–346. Springer, 2013. ISBN 978-3-642-41334-6.

[16] Viet Bach Nguyen, Vojtech Svátek, Gollam Rabby, and Ondrej Zamazal. OReCaP - towards ontology reuse via focused categorization power. In Kerry L. Taylor, Rafael S. Gonçalves, Freddy Lécué, and Jun Yan, editors, *Proceedings of the ISWC 2020 Demos and Industry Tracks: From Novel Ideas to Industrial Practice co-located with 19th International Semantic Web Conference (ISWC 2020), Globally online, November 1-6, 2020 (UTC)*, volume 2721 of *CEUR Workshop Proceedings*, pages 295–300. CEUR-WS.org, 2020. URL http://ceur-ws.org/Vol-2721/paper575.pdf.

[17] Alan L. Rector. Modularisation of domain ontologies implemented in description logics and related formalisms including OWL. In John H. Gennari, Bruce W. Porter, and Yolanda Gil, editors, *K-CAP*, pages 121–128. ACM, 2003. ISBN 1-58113-583-1.

[18] Alan L. Rector, Nick Drummond, Matthew Horridge, Jeremy Rogers, Holger Knublauch, Robert Stevens, Hai Wang, and Chris Wroe. Owl Pizzas: Practical experience of teaching OWL-DL: Common errors & common patterns. In Enrico Motta, Nigel Shadbolt, Arthur Stutt, and Nicholas Gibbins, editors, *EKAW*, volume 3257 of *Lecture Notes in Computer Science*, pages 63–81. Springer, 2004.

[19] Yuan Ren, Artemis Parvizi, Chris Mellish, Jeff Z. Pan, Kees van Deemter, and Robert Stevens. Towards competency question-driven ontology authoring. In Valentina Presutti, Claudia d'Amato, Fabien Gandon, Mathieu d'Aquin, Steffen Staab, and Anna Tordai, editors, *The Semantic Web: Trends and Challenges - 11th International Conference, ESWC 2014, Anissaras, Crete, Greece, May 25-29, 2014. Proceedings*, volume 8465 of *Lecture Notes in Computer Science*, pages 752–767. Springer, 2014. .

[20] Johann Schaible, Thomas Gottron, and Ansgar Scherp. Survey on common strategies of vocabulary reuse in linked open data modeling. In Valentina Presutti, Claudia d'Amato, Fabien Gandon, Mathieu d'Aquin, Steffen Staab, and Anna Tordai, editors, *ESWC*, volume 8465 of *Lecture Notes in Computer Science*, pages 457–472. Springer, 2014. ISBN 978-3-319-07442-9.

[21] Ioannis Stavrakantonakis, Anna Fensel, and Dieter Fensel. Linked open vocabulary ranking and terms discovery. In Anna Fensel, Amrapali Zaveri, Sebastian Hellmann, and Tassilo Pellegrini, editors, *SEMANTICS*, pages 1–8. ACM, 2016. ISBN 978-1-4503-4752-5.

[22] Robert Stevens, Phillip Lord, James Malone, and Nicolas Matentzoglu. Measuring expert performance at manually classifying domain entities under upper ontology classes. *J. Web Semant.*, 57, 2019.

[23] Vojtech Svátek, Ondrej Zamazal, and Miroslav Vacura. Categorization power of ontologies with respect to focus classes. In *Knowledge Engineering and Knowledge Management - 20th International Conference, EKAW 2016, Bologna, Italy, November 19-23, 2016, Proceedings*, pages 636–650, 2016. URL https://doi.org/10.1007/978-3-319-49004-5_41.

[24] Vojtech Svátek, Martin Homola, Jan Kluka, and Miroslav Vacura. Metamodeling-based coherence checking of OWL vocabulary background models. In Mariano Rodriguez-Muro, Simon Jupp, and Kavitha Srinivas, editors, *OWLED*, volume 1080 of *CEUR Workshop Proceedings*. CEUR-WS.org, 2013.

[25] Christoph Tempich and Raphael Volz. Towards a benchmark for semantic web reasoners - an analysis of the DAML ontology library. In York Sure and Óscar Corcho, editors, *EON*, volume 87 of *CEUR Workshop Proceedings*. CEUR-WS.org, 2003.

[26] Allan Third. "Hidden semantics": what can we learn from the names in an ontology? In *INLG*, pages 67–75, 2012.

[27] Pierre-Yves Vandenbussche, Ghislain Atemezing, María Poveda-Villalón, and Bernard Vatant. Linked open vocabularies (LOV): A gateway to reusable semantic vocabularies on the web. *Semantic Web*, 8(3):437–452, 2017.

Appendix A: Questionnaire task description from the 2016 cognitive experiment campaign

FC and category in Manchester syntax:

– Individual: ?i Types: Person
– Individual: ?i Types: bornIn some Thing

?i is an instance of a class including all objects that are in the relationship 'bornIn' to at least one object.

Is the class (bornIn some Thing) a meaningful category for categorizing objects of class Person?

Hint: the expression, e.g. (bornIn some Thing), suitable for categorization of the FC should satisfy the following conditions:

1. there may theoretically exist a higher number (i.e. not 0, 1 or just a few) FC instances satisfying the expression (absolute number of instances)
2. there may theoretically exist a higher number (i.e. not 0, 1 or just a few) FC instances *not* satisfying the expression, in turn
3. ideally (this is not a strict condition) the expression is not apparently tied to some specific subclass of the given FC, such that it would not make any sense to consider it for other subclasses of this FC.

Possible values:

– certainly (2)
– perhaps (1)
– borderline (0)
– perhaps not (-1)
– certainly not (-2)
– no judgment, since I don't understand the example (N)

Appendix B: Questionnaire task description from the 2018 cognitive experiment campaign

The focus class is from the ontology no. 162945, http://www.ontotext.com/proton/protonext

Ontology name: Proton Ontology

Ontology description:
"PROTON (PROTo ONtology) was developed in the SEKT project as a lightweight upper-level ontology, serving as a modeling basis for a number of tasks in different domains. To mention just a few applications: PROTON is meant to serve as a seed for ontology generation (new ontologies constructed by extending PROTON); it can be used for automatic entity recognition and more generally Information Extraction (IE) from text, for the sake of semantic annotation (metadata generation). PROTON was extended to cover the conceptual knowledge encoded within the most popular datasets from Linked Open Data like DBPedia, GeoNames, etc."

Involved entities:
Focus class: **Object**
Property: **owner**
Target class: **Parliament**

Proposed category:
A specific kind of Object: such that has an owner that is a Parliament

Description of entities:
Object: Objects are entities that could be claimed to exist - in some sense of existence. An object can play a certain role in some happenings. Objects could be substantially real - as the Buckingham Palace or a hardcopy book - or substantially imperceptible - for instance, an electronic document that exists only virtually, one cannot touch it.
owner: The relationship between an object and an agent who owns it.
Parliament: A legislative assembly representative at national or regional level. It can also be called Senate, etc.

It holds:
Parliament SubClassOf: PoliticalEntity
Do you understand the meaning of these entities, making use of the description above?
Possible values:

– Yes, I am quite familiar with them (3)
– Yes, roughly (2)
– I only have a pretty vague idea (1)
– Absolutely no clue (0)

Is "Object that has an owner that is a Parliament" a meaninful and reusable category for instances of the class Object?
Possible values:

– Certainly (2)
– Perhaps (1)
– Borderline (0)
– Perhaps not (-1)
– Certainly not (-2)
– No judgment, since I don't understand the example (N)