

# Проблема транзитивности в системе категорий Википедии

© А. В. Кириллович<sup>1</sup>

© О.А. Невзорова<sup>1,2</sup>

<sup>1</sup>Казанский федеральный университет, Казань, Россия

<sup>2</sup>Институт прикладной семиотики АН Республики Татарстан, Казань, Россия

alikh.kirillovich@gmail.com

onevzoro@gmail.com

**Аннотация.** Проведен анализ системы категорий Википедии. Показано, что в системе категорий происходит нарушение требования транзитивности, в результате чего статья из подкатегории некоторой категории может не относиться к основной категории. С помощью стандартных методов онтологического моделирования проанализированы причины нарушений транзитивности. Предложен подход к автоматическому устранению данных нарушений.

**Ключевые слова:** Википедия, система категорий, информационно-поисковой тезаурус, онтологическое моделирование, OntoClean, онтологическая зависимость

## Problem of Transitivity of Wikipedia Category System

© Alexander Kirillovich<sup>1</sup>

© Olga Nevzorova<sup>1,2</sup>

<sup>1</sup>Kazan Federal University, Kazan, Russia

<sup>2</sup>Research Institute of Applied Semiotics of Tatarstan Academy of Sciences, Kazan, Russia

alikh.kirillovich@gmail.com

onevzoro@gmail.com

**Abstract.** This paper analyses a violation of the transitivity principle of Wikipedia category system. Causes of the violation have been analyzed on base of ontological modeling methodologies such as OntoClean. A new approach for elimination of the violation has been proposed.

**Keywords:** Wikipedia, categorization system, thesaurus, ontology engineering, OntoClean, ontological dependence

### 1 Введение

Википедия — одно из крупнейших хранилищ информации. Данные Википедии используются в таких задачах как разрешение лексической многозначности, категоризация текста, вычисление семантической близости, машинный перевод. Для автоматической обработки информации из Википедии требуется средство ее структурирования.

Система категорий — основной инструмент структурирования информации в Википедии. Категории хранят статьи, объединенные общей темой.

Категории бывают двух видов:

- Категории-множества, например C:Cities, которая содержит статьи про конкретные города (*Москва, Казань, Лондон, Урюпинск* и т. д.).
- Категории-топики, например C:City, которая содержит статьи по городской тематике (*Городское планирование, Урбанизация, История городов, Городская культура* и т. д.).

Каждая категория может содержать подкатегории и самой находиться в родительских категориях. Таким образом, система категорий представляет собой ориентированный граф без циклов.

Категории могут группироваться с использованием мета-категорий, например C:Writers → C:Writers\_by\_nationality → C:Russian\_writers.

Согласно правилам Википедии, статья должна находиться в наиболее специфичной категории в иерархии. Например статья Anton\_Chekhov должна находиться не в категории C:Writers, а в ее вложенной подкатегории C:Russian\_novelists через цепочку: C:Writers → C:Writers\_by\_nationality → C:Russian\_writers → C:Russian\_novelists → Anton\_Chekhov. Поэтому, чтобы получить все статьи, релевантные заданной категории, требуется извлекать статьи из всех ее вложенных подкатегорий.

В связи с этим требуется, чтобы система категорий была транзитивной: статьи из вложенных подкатегорий должны быть релевантны родительской категории. Однако требование транзитивности нарушается. Так например категория *Арифметика* содержит нерелевантную ей вложенную подкатеорию *Расстрелянные колумбийцы*: C:Arithmetic → C:Ratios → C:Rates → C:Temporal\_rates → C:Acceleration → C:Force → C:Motion\_(physics) →

C:Flight → C:Ballistics → ... → C:Projectile\_weapons →  
 C:Firearms → C:People\_associated\_with\_firearms →  
 C:Shooting\_victims → ... →  
 C:Colombian\_people\_executed\_by\_firing\_squad.

Цель данной статьи — проанализировать проблему нарушения транзитивности в системе категорий Википедии и предложить подход к ее решению. Она построена следующим образом. В разделе 2 кратко перечислены другие средства структурирования данных Википедии и отмечены их ограничения. В разделе 3 система категорий исследована с помощью классических методов онтологического моделирования и выявлены причины потери транзитивности. В разделе 4 предложен подход, который исправляет систему категорий, удаляя из нее не-транзитивные цепочки. В Разделе 5 описаны направления будущей работы.

## 2 Связанные работы

Существует множество проектов по извлечению структурированных данных из Википедии [1, 2]: DBpedia [3–5], YAGO [6–8], WikiTaxonomy [9–11], WikiNet [12–14], ORA: The Natural Ontology of Wikipedia [15,16], WiBi [17], MENTA [18], BabelNet [19–20], WiSiNet [22], KOG: Kylin Ontology Generator [23], а также проектов выравнивания системы категорий и WordNet [24–28]. Однако данные проекты не являются универсальными и не применимы к категориям-топикам.

## 3 Анализ причин нарушения транзитивности

Проанализируем систему категорий с помощью методологий онтологического моделирования и выявим причины потери транзитивности.

Систему категорий можно рассмотреть в качестве тезауруса [29, 30]. Категории будут соответствовать концептам, а отношения между категорией и подкатегорией – стандартным онтологическим отношениям:

- Отношение класс → подкласс:
  - C:Cities\_in\_Europe → C:Capitals\_in\_Europe;
  - C:Software → ... → C:Operating\_systems;
  - C:Mathematical\_axioms → C:Axioms\_of\_set\_theory;
  - C:Machines → C:Engines;
  - C:Wars → ... → C:Wars\_involving\_the\_Soviet\_Union;
  - C:Fiction\_books → ... → C:Dystopian\_novels.
- Отношение класс → экземпляр:
  - C:Capitals\_in\_Europe → C:Moscow;
  - C:Intergovernmental\_organizations → C:United\_Nations;
  - C:Universities\_and\_colleges\_in\_Connecticut → C:Yale\_University;

- C:Operating\_systems → C:Unix;
- C:Fields\_of\_mathematics → C:Algebra;
- C:Axioms\_of\_set\_theory → C:Axiom\_of\_choice;
- C:Abstract\_strategy\_games → C:Chess;
- C:Engines → C:Internal\_combustion\_engine;
- C:Wars\_involving\_the\_Soviet\_Union → C:World\_War\_II;
- C:Dystopian\_novels → C:Nineteen\_Eighty-Four;
- C:Organs → C:Brain;
- C:Space\_stations → C:International\_Space\_Station.
- Отношение часть → целое:
  - C:Moscow → C:Cities\_and\_towns\_under\_jurisdiction\_of\_Moscow → C:Zelenograd;
  - C:Yale\_University → C:Yale\_University\_Library
  - C:United\_Nations → C:International\_Atomic\_Energy\_Agency;
  - C:World\_War\_II → ... → C:Attack\_on\_Pearl\_Harbor;
  - C:Central\_nervous\_system → C:Brain;
  - C:Unix → C:Network\_socket;
  - C:Internal\_combustion\_engine → C:Pistons.
- Ассоциативные отношения:
  - Наука → объект изучения:
    - C:Botany → C:Plants.
  - Агент → контрагент:
    - C:Plants → C:Herbicides;
    - C:Violence → C:Nonviolence;
    - C:Communism → C:Anti-communism.
  - Величина → инструмент для измерения:
    - C:Temperature → C:Thermometers.
  - Деятельность → агент деятельности:
    - C:Hunting → C:Hunting\_dogs;
    - C:Military → C:Military\_personnel.
  - Сырье → результат:
    - C:Grape → C:Raisins;
    - C:Petroleum → C:Petroleum\_products → Gasoline;
    - C:Textiles → C:Textile\_arts → C:Weaving.
  - Другие ассоциативные отношения:
    - C:Death → C:Death\_customs → C:Funerals;
    - C:Automobiles → C:Auto\_racing;
    - C:Books → C:Book\_arts → C:Bookbinding.

Мета-категориям соответствует так называемые «Node labels».

Задача нахождения релевантных подкатегорий для заданной категории соответствует стандартной задаче расширения поискового запроса (query expansion).

Представив систему категорий в виде тезауруса, мы применили к ней формальную методологию проверки корректности онтологий OntoClean, а также методологию построения информационно-поисковых тезаурусов [29–33]. В результате оказалось, что многие случаи нарушения транзитивности вызваны нарушениями правил построения иерархии концептов онтологии. Основными такими причинами являются:

- Неполное включение одной категории в другую:
  - Аналоговая *Киноленка* попала в категорию *Цифровые технологии*:  
C:Digital\_technology → C:Digital\_media → C:Video → C:Film\_and\_video\_technology → C:Film\_stock. Причина в том, что категория *Видео* не полностью входит в категорию *Цифровые медиа*, т.к. существует и аналоговое видео.
  - Англоязычный роман «*Лолита*» попал в категорию *Русские романы*:  
C:Russian\_novels → ... → C:Novels\_by\_Vladimir\_Nabokov → Lolita. Аналогично, категория *Новеллы Набокова* не полностью входит в категорию *Русские романы*.
  - *Японский язык* попал в категорию *Языки Кореи*:  
C:Languages\_of\_Korea → C:Buyeo\_languages → C:Japonic\_languages → C:Japanese\_language.
- Ошибки при использовании нечетких понятий:
  - *Электрические стулья* попали в категорию *Потребительские товары*:  
C:Consumer\_goods → C:Furniture → C:Chairs → C:Electric\_chairs.
- Ошибки с использованием омонимичных категорий:
  - *Музыкальные чарты* попали в *Диаграммы*:  
C:Diagrams → C:Charts → C:Record\_charts. В одном случае C:Charts использовались в значении диаграмм, а в другом – в значении музыкальных чартов.
  - *Строительство кораблей* попало в *Недвижимость*:  
C:Real\_estate → C:Construction → C:Ship\_construction.
- Использование одного понятия в разных смыслах:
  - Электронная библиотека *Lib.ru* попала в категорию *Здания*:  
C:Buildings\_and\_structures → C:Buildings\_and\_structures\_by\_type → C:Libraries → C:Digital\_libraries → Lib.ru. В одном случае *Библиотека* рассматривалась как тип здания, а в другом – как социальный институт.
  - Философское мировоззрение *Нигилизм* попало в категорию *Биология*:  
C:Biography → C:Life → C:Philosophy\_of\_life → C:Nihilism. В одном случае *Жизнь* рассматривалась как биологический процесс, а другом случае – как социальный процесс.
- *Снег* попал в *Жидкости*:  
C:Liquids → C:Water → C:Forms\_of\_water → C:Snow. В одном случае *Вода* рассматривалась как вещество вообще, а в другом – как вещество в жидком агрегатном состоянии.
- Несовместимые критерии идентичности:
  - Мусульманская святыня *Кааба* попала в категорию *Математические объекты*:  
C:Mathematical\_objects → C:Geometric\_shapes → C:Elementary\_shapes → C:Cubes → C:Cubic\_buildings → C:Kaaba. Ошибка находится в цепочке C:Cubes → C:Cubic\_buildings. Кубические здания, вообще говоря, не являются кубами, т.к. у них разные критерии идентичности. Куб — это абстрактный, вневременной, неизменный объект. Если куб изменится хотя-бы на миллиметр, то это будет уже другой куб. Кубическое же здание — это конкретный объект, существующий во времени и пространстве и сохраняющий идентичность при небольших модификациях.
  - *Бермудский треугольник* попал в категорию *Геометрические объекты*:  
C:Mathematical\_objects → C:Geometric\_shapes → C:Elementary\_shapes → C:Triangles → Bermuda\_Triangle.
- Смешение понятия и знака:
  - *Династия Габсбургов* попала в категорию *Слова и фразы*:  
C:Words\_and\_phrases → ... → C:Surnames\_of\_Swiss\_origin → C:Swiss\_families → C:Swiss\_noble\_families → C:House\_of\_Habsburg.
  - *Токсин* попал в категорию *Язык*:  
C:Language → C:Terminology → C:Biological\_terminology → Toxin. Причина ошибки в том, что токсин не является термином. Термином является слово “токсин”.
- Наследование типов от ролей:
  - *Анальгетики* попали в категорию *Запрещенные наркотики*:  
C:Illegal\_drugs → C:Morphine → C:Analgesic. Причина ошибки в том, что *Запрещенные наркотики* — это не тип, а роль, и она не должна содержать категории-типы.
  - *Бомбовые прицелы* попали в категорию *Офисные принадлежности*:  
C:Office\_equipment → C:Computers → ... → C:Analog\_computers → ... → C:Optical\_bombsights.
  - *Волчья ягода* (несъедобная) попала в категорию *Еда*:  
C:Foods → C:Fruit → C:Berries → C:Sambucus.

В следующих случаях транзитивность нарушается не вследствие ошибки, а вследствие самого принципа построения системы категорий Википедии:

- Нетранзитивность отношения класс → экземпляр:
  - Корабль «*Санта-Мария*» попал в категорию *Типы кораблей*: C:Ship\_types → ... → C:Exploration\_ships → Santa\_María\_(ship).
  - Поэма под названием «*Ода*» попала в категорию *Жанры литературы*: C:Literary\_genres → C:Poetry → ... → Ode\_(poem). Этот случай особенно опасен, т. к. существует настоящий литературный жанр с таким именем.
- Нетранзитивность ассоциативного отношения:
  - Город *Зеленоград* попал в *Европейские столицы*: C:Capitals\_in\_Europe → C:Moscow → ... → C:Zelenograd.
  - Собака *Блонди* попала в категории *Нацистских лидеров*: C:Nazi\_leaders → C:Adolf\_Hitler → Blondi.
  - *Галактическая Империя* из вымышленной вселенной «Звездных войн» попала в категорию *Североамериканские государства*: C:Northern\_American\_countries → C:United\_States → C:American\_people → ... → C:George\_Lucas → C:Star\_Wars → ... → Galactic\_Empire\_(Star\_Wars).
  - *Языки Джибути* попали в категорию *Статистика*: C:Statistics → C:Statistical\_data\_sets → C:Demographics\_by\_country → C:Demographics\_of\_Djibouti → C:Languages\_of\_Djibouti.
  - *Биологическое оружие* попало в категорию *Трудовое право*: C:Labour\_law → C:Labour\_relations → C:Occupational\_safety\_and\_health → C:Toxicology → C:Biological\_weapons.
  - *Расстрелянные колумбийцы* попали в категорию *Арифметика*: C:Arithmetic → C:Ratios → C:Rates → C:Temporal\_rates → C:Acceleration → C:Force → C:Motion\_(physics) → C:Flight → C:Ballistics → ... → C:Projectile\_weapons → C:Firearms → C:People\_associated\_with\_firearms → C:Shooting\_victims → ... → C:Colombian\_people\_executed\_by\_firing\_squad

Итак, нарушение транзитивности в системе категорий Википедии вызвано двумя группами причин. К первой группе относятся причины, связанные с нарушением правил построения иерархии концептов в онтологии. Эти нарушения могут быть устранены самими авторами Википедии. Ко второй группе относятся причины, связанные с самим принципом устройства системы категорий Википедии, главная из которых — нетранзитивность ассоциативного отношения.

## 4 Подход к устранению нарушений транзитивности

В данном разделе мы предложим метод, который устраняет не транзитивные цепочки и оставляет только транзитивные.

### 4.1. Основная идея (и ее теоретическое обоснование)

Как было показано в предыдущем разделе, одна из основных причин нарушения транзитивности состоит в том, что некоторые категории связаны с подкатегориями ассоциативным отношением, которое в общем случае не является транзитивным.

Существующие методы извлечения структурированной информации из системы категорий (например, YAGO или WikiTaxonomy) выявляют ассоциативные связи между категориями и просто устраняют их. Недостатком данных методов является то, что они исключают даже те ассоциативные связи, которые не нарушают транзитивность. В связи с этим возникает потребность в методе, который устраняет ассоциативные отношения, нарушающие транзитивность (например, *Статистика* → *Демография*), но сохраняет не нарушающие (например, *Образование* → *Учитель*). Опишем основные принципы этого метода.

Предлагаемый метод основан на подходе, который применяется в тезаурусе РуТез для установления ассоциативных отношений между концептами [30, 34–37]. В соответствии с этим подходом ассоциативное отношение между двумя концептами является транзитивным, если между концептами существует отношение онтологической зависимости.

Для формализации отношения онтологической зависимости [38–41] в РуТезе используется так называемый модально-экзистенциальный подход (modal-existential account): объект А зависит от объекта В тогда и только тогда, когда необходимо, что если существует А, то существует и В [42, 43].

Модально-экзистенциальный подход имеет ряд преимуществ, среди которых — простота и математическая строгость. Недостатком этого подхода является то, что его применение требует участия человека. В связи с этим в исходном виде он не применим для решения поставленной нами задачи.

Кроме того, модально-экзистенциальный подход был подвергнут критике с чисто онтологической точки зрения. К. Файн (Kit Fine) показал, что данный подход является слишком грубым приближением к понятию онтологической зависимости и имеет ряд контрпримеров [44, 45]. В качестве альтернативы Файн предложил эссенциальный подход (essential account). Согласно этому подходу, А зависит от В, если А является неустранимой компонентой сущности В. При этом сущность объекта (essence) понимается как набор утверждений, истинных в силу идентичности данного объекта. Эти утверждения, в свою очередь, образуют реальное определение объекта (real definition) [44, 45].

Понятие *быть истинным в силу идентичности объекта* (*to be true in virtue of the identity of*) требует пояснения. В соответствии с интерпретацией самого Файна, основанной на позиции мета-онтологического реализма, данное понятие выражает объективно существующее, примитивное и неопределимое отношение между утверждением и объектом. Мы предлагаем другую интерпретацию, основанную на подходе Э.Л. Томассон (Amie L. Thomasson). Согласно Томассон, условиями существования и идентичности некоторого объекта А являются условия применения/повторного применения термина “А” (application/co-application conditions) [46-47]. В соответствии с нашей интерпретацией, утверждение является истинным в силу идентичности объекта А, если оно является частью условий применения/повторного применения термина “А”. И соответственно, объект А онтологически зависит от объекта В, если условия применения/повторного применения термина “А” включают в себя условия применения/повторного применения термина “В”. Наша интерпретация показывает, что эссенциальный подход не требует принятия конкретных философских предпосылок.

Таким образом, используя эссенциальный подход, мы получили следующий критерий определения онтологической зависимости: X онтологически зависит от Y, если Y неустранимым образом входит в определение X. Данный критерий гораздо лучше подходит для автоматического применения.

В качестве аппроксимации определения объекта, соответствующего той или иной категории, мы брали аннотации главной статьи данной категории и этой статьи на других языках. Факт вхождения объекта в определение другого объекта моделировался как наличие гиперссылки между определениями.

#### 4.2. Опыт реализации

Экспериментальная реализация предложенного подхода выглядит следующим образом:

- Определяем, является ли отношение между категорией и ее подкатегорией ассоциативным. Полагаем, что отношение является ассоциативным, если в нем участвует категория-топик. Тип категории определяем с помощью метода из проекта WikiTaxonomy [11].
- Если отношение является ассоциативным, то с помощью описанного выше критерия проверяем отношение онтологической зависимости между подкатегорией и категорией. Если зависимость имеется, то сохраняем отношение между категориями, если не имеется, то удаляем.
- Если отношение является не ассоциативным, а таксономическим, то используем уже существующий ресурс YAGO, содержащий очищенные таксономические отношения. В случае, если отношение присутствует в YAGO, сохраняем его и удаляем в противоположном случае.

#### 4.3. Оценка

Мы провели предварительное оценивание нашего метода на категории C:Mathematics. Выбор этой категории был обусловлен решением практической задачи по связыванию онтологии OntoMath<sup>Pro</sup> [48-50] с DBpedia [3-5]. Для этого с помощью данного метода мы исключили из этой категории предположительно нерелевантные ей подкатегории. Список удаленных и оставленных подкатегорий был передан для ручной оценки. Задача ассессора состояла в том, чтобы оценить, действительно ли оставленные категории релевантны основной категории и действительно ли удаленные – не релевантны. Результат оценивания представлен в Таблице 1.

**Таблица 1** Результат предварительной оценки предложенного метода на категории C:Mathematics

Total	4281
True positives	2136
True negatives	650
False positives	1010
False negatives	485
Recall	0,814956
Precision	0,678957
F1 score	0,740766

#### 5 Заключение

Мы проанализировали причины нарушения транзитивности в системе категорий Википедии и предложили подход к их устранению.

В дальнейшем мы планируем доработать данный подход. В частности, предполагается извлекать определения из других страниц категории, а также использовать контекст ссылки внутри определения.

Мы планируем применить разработанный подход для извлечения тезауруса из системы категорий Википедии. В отличие от тезаурусов, построенных вручную (таких как RuТез или WordNet), данный тезаурус будет содержать не только общие понятия, но и большое число именованных сущностей. В отличие от других тезаурусов, автоматически извлеченных их системы категорий Википедии, данный тезаурус будет содержать не только иерархические отношения, но и подмножество ассоциативных отношений, сохраняющих свойство транзитивности.

Извлеченный тезаурус будет интегрирован в RuThes Cloud, многоуровневый ресурс Лингвистических открытых связанных данных (LLOD) для русского языка и языков народов России [51].

**Благодарности.** А. Кириллович благодарит проф. Сорена Ауэра (Sören Auer) из Боннского университета (Universität Bonn) за доброжелательность и ценные советы на ранней стадии работы над данным проектом.

Работа выполнена за счет средств субсидии, выделенной Казанскому федеральному университету для выполнения гос. задания в сфере научной деятельности, проект 1.2368.2017/ПЧ, а также при поддержке правительственной программы повышения конкурентоспособности Казанского федерального университета.

## Литература

- [1] Medelyan, O., et al. Mining meaning from Wikipedia. *International Journal of Human-Computer Studies*, 67 (9), pp. 716-754. Academic Press (2009). doi:10.1016/j.ijhcs.2009.05.004
- [2] Hovy, E., Navigli, R., Ponzetto, S.P. Collaboratively built semi-structured content and Artificial Intelligence: The story so far. *Artificial Intelligence*, 194, pp. 2–27. Elsevier (2013). doi:10.1016/j.artint.2012.10.002
- [3] Auer, S., et al. DBpedia: A Nucleus for a Web of Open Data. In: Aberer K. et al. (eds.) 6th International Semantic Web Conference, 2nd Asian Semantic Web Conference (ISWC 2007 + ASWC 2007). *The Semantic Web. LNCS*, 4825, pp. 722-735. Springer (2007). doi:10.1007/978-3-540-76298-0\_52
- [4] Bizer, C., et al. DBpedia: A Crystallization Point for the Web of Data. *Journal of Web Semantics*, 7 (3), pp. 154–165. Elsevier (2009). doi:10.1016/j.websem.2009.07.002
- [5] Lehmann, J., et al. DBpedia: A Large-scale, Multilingual Knowledge Base Extracted from Wikipedia. *Semantic Web Journal*, 6 (2), pp. 167-195. IOS Press (2015). doi:10.3233/SW-140134
- [6] Suchanek, F. M., Kasneci, G., Weikum, G. Yago: a Core of Semantic Knowledge. In: 16th international conference on World Wide Web (WWW 2007), pp. 697-706 (2007). doi:10.1145/1242572.1242667
- [7] Hoffart, J., et al. G. YAGO2: A spatially and temporally enhanced knowledge base from Wikipedia. *Artificial Intelligence*, 194, pp. 28–61. Elsevier (2013). doi:10.1016/j.artint.2012.06.001
- [8] Mahdisoltani, F., Biega, J., Suchanek, F.M. YAGO3: A Knowledge Base from Multilingual Wikipedias. In: 7th Biennial Conference on Innovative Data Systems Research (CIDR 2015) (2015)
- [9] Ponzetto, S.P., Strube, M. Deriving a Large Scale Taxonomy from Wikipedia. In: 22nd National Conference on Artificial Intelligence (AAAI 2007), pp. 1440-1445. AAAI Press (2007)
- [10] Ponzetto, S.P., Strube, M. Taxonomy induction based on a collaboratively built knowledge repository. *Artificial Intelligence*, 175 (9–10), pp. 1737–1756. Elsevier (2011). doi:10.1016/j.artint.2011.01.003
- [11] Zirn, C., Nastase, V., Strube, M. Distinguishing between Instances and Classes in the Wikipedia Taxonomy. In: Zirn C., et al. (eds.) 5th European Semantic Web Conference (ESWC 2008). *LNCS*, 5021, pp. 376-387. Springer (2008). doi:10.1007/978-3-540-68234-9\_29
- [12] Nastase, V., Strube, M. Decoding Wikipedia Categories for Knowledge Acquisition. In: Cohn A. (ed.) 23rd National Conference on Artificial Intelligence (AAAI 2008), vol. 2, pp. 1219-1224. AAAI Press (2008)
- [13] Nastase, V., et al. WikiNet: A Very Large Scale Multi-Lingual Concept Network. In: Calzolari N., et al. (eds.) 7th International Conference on Language Resources and Evaluation (LREC 2010), pp. 1015-1022 (2010)
- [14] Nastase, V., Strube, M. Transforming Wikipedia into a large scale multilingual concept network. *Artificial Intelligence*, 194, pp. 62–85. Elsevier (2013). doi:10.1016/j.artint.2012.06.008
- [15] Gangemi, A., et al. Automatic Typing of DBpedia Entities. In: Cudré-Mauroux P. et al. (eds.) 11th International Semantic Web Conference (ISWC 2012). *LNCS*, 7649, pp. 65-81. Springer (2012). doi:10.1007/978-3-642-35176-1\_5
- [16] Nuzzolese, A.G., et al. Towards the Natural Ontology of Wikipedia. In: Blomqvist E., Groza T. (eds.) Proceedings of the ISWC 2013 Posters & Demonstrations Track. *CEUR Workshop Proceedings*, vol. 1035, pp. 273-276 (2013)
- [17] Flati, T., et al. Two Is Bigger (and Better) Than One: the Wikipedia Bitaxonomy Project. In: Toutanova K., Wu H. (eds.) 52nd Annual Meeting of the Association for Computational Linguistics (ACL 2014), vol. 1, pp. 945-955. ACM (2014). doi:10.3115/v1/P14-1089
- [18] de Melo, G., Weikum, G. MENTA: Inducing Multilingual Taxonomies from Wikipedia. In: 19th ACM international conference on Information and knowledge management (CIKM '10), pp. 1099-1108. ACM (2010). doi:10.1145/1871437.1871577
- [19] Navigli, R., Ponzetto, S.P. BabelNet: Building a Very Large Multilingual Semantic Network. In: 48th Annual Meeting of the Association for Computational Linguistics (ACL 2010), pp. 216-225. ACL (2010)
- [20] Navigli, R., Ponzetto, S.P. BabelNet: The automatic construction, evaluation and application of a wide-coverage multilingual semantic network. *Artificial Intelligence*, 193, pp. 217–250. Elsevier (2012). doi:10.1016/j.artint.2012.07.001
- [21] Moro, A., Navigli, R. WiSeNet: Building a Wikipedia-based Semantic Network with Ontologized Relations. In: 21st ACM international conference on Information and knowledge management (CIKM 2012), pp. 1672-1676. ACM (2012). doi:10.1145/2396761.2398495
- [22] Wu, F., Weld, D.S. Automatically Refining the Wikipedia Infobox Ontology. In: 17th international conference on World Wide Web (WWW 2008), pp. 635-644. ACM (2008). doi:10.1145/1367497.1367583
- [23] Ruiz-Casado, M., Alfonseca, E., Castells, P. Automatic Assignment of Wikipedia Encyclopedic Entries to WordNet Synsets. In: Szczepaniak P.S., et al. (eds.) 3rd International Conference on Advances in Web Intelligence (AWIC 2005). *LNCS*, 3528, pp. 380-386. Springer (2005). doi:10.1007/11495772\_59
- [24] Toral, A., Muñoz, R., Monachini, M. Named Entity WordNet. In: 6th Conference on Language Resources and Evaluation (LREC 2008), pp. 741-747 (2008)

- [25] Niemann, E., Gurevych, I. The people's web meets linguistic knowledge: automatic sense alignment of Wikipedia and WordNet. In: 9th International Conference on Computational Semantics (IWCS 2011), pp. 205–214. ACL (2011)
- [26] Ponzetto, S.P., Navigli, R. Large-Scale Taxonomy Mapping for Restructuring and Integrating Wikipedia. In: 21st International Joint Conference on Artificial Intelligence (IJCAI 2009), pp. 2083–2088. Morgan Kaufmann Publishers (2009)
- [27] Gella, S., Strapparava, C., Nastase, V. Mapping WordNet Domains, WordNet Topics and Wikipedia Categories to Generate Multilingual Domain Specific Resources. In: Calzolari N., et al. (eds.) 9th International Conference on Language Resources and Evaluation (LREC 2014), pp. 1117–1121. ELRA (2014)
- [28] Titze, G., et al. DBpedia Domains: augmenting DBpedia with domain information. In: Calzolari N., et al. (eds.) 9th International Conference on Language Resources and Evaluation (LREC 2014), pp. 1438–1442. ELRA (2014)
- [29] ANSI-NISO Z39.19-2005
- [30] Loukachevitch, N. Thesauri in Information Retrieval Tasks. Moscow University Press (2011)
- [31] Guarino, N., Welty, C. An Overview of OntoClean. In: Staab S., Studer R. (eds.) Handbook on Ontologies. 2nd edition. Springer (2009). doi:10.1007/978-3-540-92673-3\_9
- [32] Guarino, N., Welty, C. A Formal Ontology of Properties. In: Dieng R., Corby O. (eds.) Knowledge Engineering and Knowledge Management Methods, Models, and Tools (EKAW 2000). LNCS, 1937, pp. 97–112. Springer (2000). doi:10.1007/3-540-39967-4\_8
- [33] Guizzardi, G. Ontological foundations for structural conceptual models. CTIT (2005)
- [34] Loukachevitch, N., Dobrov, B. RuThes Linguistic Ontology vs. Russian Wordnets. In: Orav H., Fellbaum C., Vossen P. (eds.) 7th Conference on Global WordNet (GWC 2014), pp. 154–162. University of Tartu Press (2014)
- [35] Loukachevitch, N., Dobrov, B., Chetviorkin, I. RuThes-Lite, a Publicly Available Version of Thesauri of Russian Language RuThes. In: Computational Linguistics and Intellectual Technologies: Papers from the Annual International Conference “Dialogue”, pp. 340–349. RGGU (2014)
- [36] Loukachevitch, N., Dobrov, B. Development of Ontologies with Minimal Set of Conceptual Relations. In: Lino, M.T., et al. (eds.) 4th International Conference on Language Resources and Evaluation (LREC'04), pp. 1889–1892. ELRA (2004)
- [37] Loukachevitch, N., Dobrov, B. Ontological Types of Associative Relations in Information-Retrieval Thesauri and Automatic Query Expansion. In: Oltramari A. et al. (eds.) Ontologies and Lexical Resources in Distributed Environments (Ontolex 2004), pp. 24–29 (2004)
- [38] Tahko, T.E., Lowe, E.J. Ontological Dependence. In: Zalta E.N. (ed.) The Stanford Encyclopedia of Philosophy (2016). <http://plato.stanford.edu/entries/dependence-ontological/>
- [39] Correia, F. Ontological Dependence. *Philosophy Compass*, 3 (5), pp. 1013–1032. Wiley (2008). doi:10.1111/j.1747-9991.2008.00170.x
- [40] Koslicki, K. Varieties of Ontological Dependence. In: Correia F., Schnieder B. (eds.) *Metaphysical Grounding: Understanding the Structure of Reality*, pp. 186–213. Cambridge University Press (2012)
- [41] Koslicki, K. Ontological Dependence: An Opinionated Survey. In: Hoeltje M., et al. (eds.) *Varieties of Dependence*, pp. 31–64. Philosophia Verlag (2013)
- [42] Simons, P. *Parts: A Study in Ontology*. Clarendon Press (1987). Ch. 8. Ontological Dependence. doi:10.1093/acprof:oso/9780199241460.001.0001
- [43] Thomasson, A.L. *Fiction and Metaphysics*. Cambridge University Press. Chapter 2. The Nature and Varieties of Existential Dependence, pp. 24–34 (1999)
- [44] Fine, K. Essence and Modality. *Philosophical Perspectives*, 8, pp. 1–16. Ridgeview Publishing Company (1994). doi:10.2307/2214160
- [45] Fine, K. Ontological Dependence. *Proceedings of the Aristotelian Society*, 95, pp. 269–290. Wiley (1995)
- [46] Thomasson, A.L. Existence Questions. *Philosophical Studies*, 141(1), pp. 63–78. Springer (2008). doi:10.1007/s11098-008-9263-8
- [47] Thomasson, A.L. Answerable and Unanswerable Questions. In: Chalmers D.J., et al. (eds.) *Metametaphysics: New Essays on the Foundations of Ontology*. Oxford University Press (2009)
- [48] Nevzorova, O., et al. OntoMath<sup>PRO</sup> Ontology: a Linked Data Hub for Mathematics. In: Klinov P., Mouromstev D. (eds.) 5th International Conference on Knowledge Engineering and Semantic Web (KESW 2014). CCIS, vol. 468, pp. 105–119. Springer, Heidelberg (2014). doi:10.1007/978-3-319-11716-4\_9
- [49] Elizarov, A.M., et al. Mathematical Knowledge Representation: Semantic Models and Formalisms. *Lobachevskii J. Math.* 35(4), pp. 348–354. Pleiades Publishing (2014). doi:10.1134/S1995080214040143
- [50] Elizarov, A., et al. Digital Ecosystem OntoMath: Mathematical Knowledge Analytics and Management. In: Kalinichenko L., et al. (eds.) DAMDID/RCDL 2016. CCIS, vol. 706, pp. 33–46. Springer, Cham (2017). doi:10.1007/978-3-319-57135-5\_3
- [51] Kirillovich, A., et al. RuThes Cloud: Towards a Multilevel Linguistic Linked Open Data Resource for Russian. In: Rózewski P., Lange C. (eds.) KESW 2017. CCIS, vol. 786, pp. 38–52. Springer, Cham (2017). doi:10.1007/978-3-319-69548-8\_4