

Министерство науки и высшего образования Российской Федерации  
НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ  
ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ (НИ ТГУ)

САЕ “Институт человека цифровой эпохи”

УТВЕРЖДАЮ:  
Руководитель ОПОП:

 3.И. Резанова

« 31 » августа 20 22 г.

Рабочая программа дисциплины

**Формальная семантика. Онтологии и тезаурусы**

по направлению подготовки

**45.04.03 Фундаментальная и прикладная лингвистика**

Направленность (профиль) подготовки :

**Компьютерная и когнитивная лингвистика**

Форма обучения

**Очная**

Квалификация

**Магистр**


Год приема

**2022**

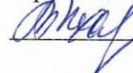
Код дисциплины в учебном плане: Б1.В.ДВ.1.1.4

СОГЛАСОВАНО:

Руководитель ОПОП

 3.И. Резанова

Председатель УМК

 Ю.А. Тихомирова

Томск – 2022

## **1. Цель и планируемые результаты освоения дисциплины**

Целью освоения дисциплины является формирование следующих компетенций:

ПК-3 Способность разрабатывать системы автоматической обработки звучащей речи и письменного текста на естественном языке, лингвистические компоненты электронных ресурсов и интеллектуальных электронных систем (лингвистические корпуса, словари, онтологии, базы данных)

ПК-4 Способность разрабатывать проекты прикладной направленности в области когнитивной и компьютерной лингвистики с применением современных технических средств и информационных технологий, в том числе в области искусственного интеллекта

Результатами освоения дисциплины являются следующие индикаторы достижения компетенций:

ИПК-4.3 Обеспечивает выполнение проекта в области когнитивной и компьютерной лингвистики с применением современных технических средств и информационных технологий, в том числе в области искусственного интеллекта, в соответствии с установленными целями, сроками и затратами..

ИПК-4.2 Разрабатывает программу действий по решению задач проекта в области когнитивной и компьютерной лингвистики с учетом имеющихся технических средств и информационных технологий, в том числе в области искусственного интеллекта..

ИПК-3.3 Разрабатывает лингвистические компоненты интеллектуальных информационных систем (онтологии, базы данных).

## **2. Задачи освоения дисциплины**

– получение студентом знаний об основных теоретических и прикладных направлениях науки в области компьютерной лингвистики, формальной семантики, построении тезаурусов и онтологий с использованием современных языков разметки.

## **3. Место дисциплины в структуре образовательной программы**

Дисциплина относится к части образовательной программы, формируемой участниками образовательных отношений, предлагается обучающимся на выбор. Дисциплина входит в модуль Компьютерная лингвистика.

## **4. Семестр(ы) освоения и форма(ы) промежуточной аттестации по дисциплине**

Третий семестр, зачет с оценкой

## **5. Входные требования для освоения дисциплины**

Для успешного освоения дисциплины требуются результаты обучения по следующим дисциплинам: Введение в анализ естественного языка (NLP).

## **6. Язык реализации**

Русский

## **7. Объем дисциплины**

Общая трудоемкость дисциплины составляет 4 з.е., 144 часов, из которых:

-лекции: 8 ч.

-практические занятия: 30 ч.

Объем самостоятельной работы студента определен учебным планом.

## 8. Содержание дисциплины, структурированное по темам

Тема 1. Введение в проблематику онтологического моделирования. Формальная семантика: определения, место в лингвистике. Основные подходы. Семантические примитивы, декомпозиция существительных и глаголов. Онтология, семантический веб, RDF, RDFS, OWL и дескриптивные логики

Тема 2. Грамматика Монтегю. Грамматика Монтегю: фрагмент грамматики предикатов; логические операторы. Грамматика Монтегю: теория типов; лямбда-оператор; квантификация

Тема 3. Язык OWL. OWL: введение, общая концепция. Создание онтологии. Классы, иерархии. OWL: Protege 5. Свойства и их типы. Описание и определение классов. OWL: Protege 5. Ограничения. Аксиомы и инференции. Заключительные замечания. Контроль онтологии, отслеживание ошибок

Тема 4. Онтологии и тезаурусы: сходства и различия. Системы Art & Architecture Thesaurus® Online, The National Agricultural Library's Agricultural Thesaurus, STW Thesaurus for Economics. Изучение ГОСТов и ISO-стандартов

Тема 5. Лучшие практики в индустрии. Конференции по онтологическому моделированию: презентация доклада

Тема 6. Собственная онтология: отладка и презентация. Отладка онтологии. Презентация онтологии

## 9. Текущий контроль по дисциплине

Текущий контроль по дисциплине проводится путем контроля посещаемости, проведения контрольных работ, выполнения домашних заданий и фиксируется в форме контрольной точки не менее одного раза в семестр.

Типовые задания для проведения текущего контроля успеваемости по дисциплине

1. Напишите деривации и перевод следующих предложений, используя формат записи языка Lp, представленный в пособии.

1. Этель отравила кота.
2. Студент дал торт лектору.
3. Пес был бешеный.

2. Переведите следующие предложения на Lp и проверьте их истинность или ложность относительно M1, представленной в пособии.

1. Этель была (тем самым) гольфистом.
2. Студент был (тем самым) певцом.

3. Оцените истинность или ложность выражений 1-16 относительно множеств A-F.

$$A = \{1, 3, 5, 7\}$$

$$D = \{7, 3, 1, 5\}$$

$$B = \{a, c, e, g, i\}$$

$$E = \{e, a, i, c, i, g\}$$

$$C = \{1, a, 3, 5, e, 7\}$$

$$F = \{\{1\}, \{1,3\}, \{1,3,5\}, \{1,3,5,7\}\}$$

1.  $A = B$
2.  $e \in C$
3.  $D \in F$
4.  $D \subset C$
5.  $B \subset E$
6.  $C \cap B = \emptyset$
7.  $E = B$
8.  $a \in D$
9.  $A \in C$
10.  $D \subset F$
11.  $A \cap D = A$
12.  $D =$  первые четыре положительных нечетных числа
13.  $\{1\} \in A$
14.  $\{1,3\} \supseteq A$
15.  $E \subseteq B$
16.  $A \cup D = D$

#### Упражнение 1: Сценарий Курсов, Домен и Диапазон

Примените операцию Домен и Диапазон к Свойству, чтобы выразить следующие аксиомы:

1. Свойство записанна имеет домен Студент и диапазон Модуль.
2. Свойство сдал имеет домен Студент и диапазон Модуль.
3. Свойство вноситвкладВ имеет домен Модуль и диапазон Диплом.
4. Свойство преподает имеет домен Лектор и диапазон Модуль.

#### Упражнение 2: Сценарий Курсов. Слияние Классов.

В каждом из следующих случае:

- 1) Нарисуйте две диаграммы, применив необходимую операцию.
- 2) Определите зоны в одной из двух диаграмм (диаграммы, в которой будут кривые из другой диаграммы).
- 3) Выберите кривую, которая должна быть скопирована и определите, должна ли быть каждая зона внутри, снаружи или разделена кривой.
- 4) Скопируйте выбранную кривую соответствующим образом.
- 5) Если необходимо скопировать две кривые, то определите, каким образом скопировать вторую кривую: следуйте тому же алгоритмическому процессу, что и для первой кривой. Допускается интуитивное решение данной задачи.

Аксиомы:

1. Испытание включает Экзамен. Испытание включает Тест.
2. Человека включает Выпускник. Студент и Выпускник не пересекаются.
3. Человек включает Студент. Студент и Ученик эквивалентны.
4. ДипломСОтличием и ДипломБезОтличия не пересекаются. Диплом и Курс эквивалентны.
5. Лектор и Ученый эквивалентны. Ученик и Студент эквивалентны.
6. Человек включает Выпускник. Диплом включает ДипломБезОтличия.

Примените те же принципы, чтобы слить три аксиомы: Человек включает Выпускник; Человек включает Студент, Диплом включает ДипломБезОтличия.

Источник заданий: Stapleton G., Howse J., Delaney A. *Ontology Engineering with Diagrams: A Practical User Guide*. URL: [www.ontologyengineering.org](http://www.ontologyengineering.org). Перевод наш.

## **10. Порядок проведения и критерии оценивания промежуточной аттестации**

Зачет с оценкой в третьем семестре проводится в письменной форме и в форме презентации. Промежуточная аттестация и ее оценивание осуществляются следующим образом.

Оценка «удовлетворительно» выставляется, если онтология отвечает 60 процентам критериев, представленных ниже.

Оценка «хорошо» выставляется, если онтология отвечает 80 процентам критериев, представленных ниже.

Оценка «отлично» выставляется, если онтология отвечает более 80 процентам критериев, представленных ниже.

### Критерии оценки.

#### Охват и содержание

1. Онтология имеет сбалансированный охват моделируемой предметной области. Широта и глубина охвата онтологии должна быть примерно одинаковой во всех частях модели.

2. Использовать существующие структуры, онтологии и словари в максимально возможной степени. Подразумевается подключение схем БД, таксономий, контролируемых словарей, спецификаций и неформальных списков, тезаурусов. Использовать уже существующие проверенные дизайнерские решения.

3. Онтология встраивается в соответствующий контекст и обеспечивает возможность взаимодействия с другими онтологиями.

4. Онтология определяет четкие предикаты (свойства, отношения, атрибуты, грани, слоты), в том числе точные дефиниции. Онтология выстраивается от общей логической таксономии к более конкретным предикатам.

5. Отношения внутри онтологии должны быть когерентными, т.е. логичными и последовательными. Референциальная структура должна быть открыта к расширению и отвечать здравому смыслу.

#### Структура и дизайн

1. Онтология основывается на легкой доменной онтологии, использующей небольшое количество предикатов.

2. Онтология структурно разделяет доменные понятия (классы, субъектные концепты) от отдельных элементов (именованных сущностей, индивидов).

3. Онтология построена накопительно, от малого количества данных и связей к большему.

4. Онтология использует модульную структуру (для крупных проектов).

5. Онтология использует домены и диапазоны свойств, что в дальнейшем позволяет простраивать аксиомы и использовать системы логического вывода.

6. Онтология использует ограничения на свойства и непересекающиеся классы.

7. Онтология написана на машинно-обрабатываемом языке.

8. Онтология широко использует аннотации.

#### Названия и вокабуляр

1. Все понятия названы существительными в единственном числе с использованием CamelCase.
2. Все свойства обозначены с глаголом в виде имеетСвойство. Используется mixedCase нотация.
3. Используются понятные описательные префиксы и суффиксы (например, свойства, связанные с языком могут иметь префикс Язык.имеетСвойство).
4. Онтология использует обратные свойства там, где это уместно.
5. Все понятия и свойства имеют определения, т.е. онтология имеет семантику.
6. Онтология снабжена предпочтительными ярлыками для удобочитаемости и большей упорядоченности онтологии.

#### Документация

1. Онтология документирована, имеет встроенные комментарии.
2. Задokumentирован вокабуляр онтологии и ее методология.
3. Онтология снабжена сопутствующей вики.
4. Документация содержит полное графическое представление онтологии.

#### Тестирование

1. Онтология должна проходить тестирование с помощью системы логического вывода.
2. Онтология использует необходимые и достаточные условия.

Источник: [http://wiki.opensemanticframework.org/index.php/Ontology\\_Best\\_Practices](http://wiki.opensemanticframework.org/index.php/Ontology_Best_Practices).  
Перевод наш.

Первая часть работы проверяет компетенции ИПК-4.3, ИПК-4.2, ИПК-3.3. Вторая часть работы проверяет компетенции ИПК-4.3, ИПК-4.2, ИПК-3.3.

### **11. Учебно-методическое обеспечение**

а) Электронный учебный курс по дисциплине в электронном университете «Moodle» - <https://moodle.tsu.ru/course/view.php?id=14708>

б) Оценочные материалы текущего контроля и промежуточной аттестации по дисциплине представлены в курсе «Moodle»

в) План практических занятий по дисциплине соответствует п. 8 (лекции не предусмотрены).

г) Методические указания по организации самостоятельной работы студентов.

Самостоятельная работа студентов включает чтение текстов из списка литературы, подготовку к лекциям по технологии перевернутого класса. Подробные методические указания представлены в курсе «Moodle» - <https://moodle.tsu.ru/mod/page/view.php?id=197144>.

### **12. Перечень учебной литературы и ресурсов сети Интернет**

а) основная литература:

1. Allemang D., Hendler J., F. Gandon Semantic web for the working ontologist: effective modeling in RDFS and OWL. – Elsevier, 2020.
2. Cann R. Formal semantics: an introduction. – Cambridge University Press, 1993.

3. Horridge M. A Practical Guide To Building OWL Ontologies Using Protege 4 and CO-ODE Tools. – The University Of Manchester, 2011. URL: <http://owl.cs.manchester.ac.uk/publications/talks-and-tutorials/protg-owl-tutorial/>.
4. Kilgarriff A., Yallop C. What's in a Thesaurus? //LREC. – 2000. URL: <https://pdfs.semanticscholar.org/8ec7/f48b58e9d4adc0fb807b0426004fdf1d4b0d.pdf>
5. Stapleton G., Howse J., Delaney A. Ontology Engineering with Diagrams: A Practical User Guide. URL: [www.ontologyengineering.org](http://www.ontologyengineering.org).
6. Aloni M., Dekker P. (ed.). The Cambridge Handbook of Formal Semantics. – Cambridge University Press, 2016.

б) дополнительная литература:

7. Noy N.F., McGuinness D.L. Ontology Development 101: A Guide to Creating Your First Ontology. – Stanford University, 2014. URL: <https://protegewiki.stanford.edu/wiki/Ontology101>.
8. Portner P. What is Meaning? Fundamentals of Formal Semantics. – Oxford: Blackwell, 2005.
9. Staab S., Studer R. (ed.). Handbook on ontologies. – Springer Science & Business Media, 2010.
10. Von Heusinger K., Maienborn C., Portner P. Semantics: An international handbook of natural language meaning. – Mouton de Gruyter, 2012. – 945 p.
11. Андреев А.В., Митрофанова О.А., Соколов К.В. Введение в формальную семантику: учеб. пособие. – СПб.: СПбГУ. РИО. Филологический факультет, 2014. – 88 с.
12. Бах Э. Неформальные лекции по формальной семантике: Пер. с англ. / Под ред. О. Л. Митрофановой, О. В. Митрениной; Предисл. Б. Парти. – М.: Книжный дом «ЛИБРОКОМ», 2010. – 224 с.
13. Логическая семантика: перспективы для философии языка и эпистемологии: Сборник научных статей, посвященных юбилею Е.Д. Смирновой / Отв. ред. Е.Г. Драгалина-Черная и Д.В. Зайцев. – Москва: Креативная экономика, 2011. – 328 с.
14. Фреге Готтлоб. Логика и логическая семантика: Сборник трудов/Пер. с нем. В. В. Бирюкова, под ред. З. А. Кузичевой: Учебное пособие для студентов вузов. – М.: Аспект Пресс, 2000. – 512 с.

в) ресурсы сети Интернет:

1. Art & Architecture Thesaurus® Online [Электронный ресурс]. – URL: <http://zbw.eu/stw/versions/latest/>
2. BabelNet [Электронный ресурс]. – URL: <http://babelnet.org/>
3. Elibrary.ru: научная электронная библиотека [Электронный ресурс]. URL: [http://elibrary.ru/project\\_risc.asp](http://elibrary.ru/project_risc.asp) (01.09.2016).
4. Ontopia [Электронный ресурс]. – URL: <http://www.ontopia.net/>
5. Protege Ontology Library [Электронный ресурс]. – URL: [https://protegewiki.stanford.edu/wiki/Protege\\_Ontology\\_Library](https://protegewiki.stanford.edu/wiki/Protege_Ontology_Library)
6. Protégé OWL Tutorial [Электронный ресурс]. – URL: <http://owl.cs.manchester.ac.uk/publications/talks-and-tutorials/protg-owl-tutorial/>
7. Scopus: database [Электронный ресурс]. URL: <https://www.scopus.com/> (01.09.2016).
8. STW Thesaurus for Economics [Электронный ресурс]. – URL: <http://www.getty.edu/research/tools/vocabularies/aat/>
9. The National Agricultural Library's Agricultural Thesaurus [Электронный ресурс]. – URL: <https://agclass.nal.usda.gov/agt.shtml>
10. VOWL: Visual Notation for OWL Ontologies [Электронный ресурс]. – URL: <http://vowl.visualdataweb.org/v2/>

11. Web of Science: database [Электронный ресурс]. – URL: <http://login.webofknowledge.com/> (01.09.2016).

12. ЭБС Ebrary [Электронный ресурс]. – URL: <http://site.ebrary.com/lib/tomskuniv/>

13. Перечень информационных технологий

а) лицензионное и свободно распространяемое программное обеспечение:

– Microsoft Office 2019 Russian: пакет программ. Включает приложения: MS Office Word, MS Office Excel, MS Office PowerPoint, MS Office OneNote, MS Office Publisher, MS Outlook, MS Office Web Apps (Word Excel MS PowerPoint Outlook);

– публично доступные облачные технологии (Google Docs, Яндекс диск и т.п.).

б) информационные справочные системы:

– Электронный каталог Научной библиотеки ТГУ – <http://chamo.lib.tsu.ru/search/query?locale=ru&theme=system>

– Электронная библиотека (репозиторий) ТГУ – <http://vital.lib.tsu.ru/vital/access/manager/Index>

– ЭБС Лань – <http://e.lanbook.com/>

– ЭБС Консультант студента – <http://www.studentlibrary.ru/>

– Образовательная платформа Юрайт – <https://urait.ru/>

– ЭБС ZNANIUM.com – <https://znanium.com/>

– ЭБС IPRbooks – <http://www.iprbookshop.ru/>

#### **14. Материально-техническое обеспечение**

Аудитории для проведения занятий лекционного типа.

Аудитории для проведения занятий семинарского типа, индивидуальных и групповых консультаций, текущего контроля и промежуточной аттестации.

Помещения для самостоятельной работы, оснащенные компьютерной техникой и доступом к сети Интернет, в электронную информационно-образовательную среду и к информационным справочным системам.

#### **15. Информация о разработчиках**

Шиляев Константин Сергеевич, к. филол. н., доцент, кафедра общей, компьютерной и когнитивной лингвистики