

# Обзор докладов на VI Всероссийской конференции ЗОНТ-2017, 2-6 ноября 2017, Новосибирск, Институт Математики им. С. Л. Соболева СО РАН

к.т.н. Черкашин Е.А.

лаб. 4.1. Комплексных информационных систем

ИДСТУ СО РАН

# Общая информация

Конференция посвящена математическим методам представления данных, извлечения знаний и построения теорий предметных областей, анализу формальных понятий, методам извлечения информации из текстов естественного языка.

1. Доклады опубликованы в сборнике РИНЦ.
2. На конференции функционировала одна секция.
3. Зарегистрировано 108 участников.

# Общая информация

Тематика Конференции отражает основные стадии процесса познания:

- **Обнаружение закономерностей и извлечение знаний**, скрытых в структурированных и неструктурированных данных. Машинное обучение. Распознавание образов. Прогнозирование. Индуктивный вывод.

- **Систематизация знаний**. Инженерия знаний. Управление знаниями.

Разработка онтологий предметных областей: технологии создания и применения онтологий.

- **Построение теорий предметных областей**. Анализ формальных понятий.

Логическая семантика естественного языка. Нечеткие логики.

# Пленарные доклады

# Тезаурус RuWordNet: структура и современное состояние

Лукашевич Наталья Валентиновна

1. Русский тезаурус - это прежде всего онтология;
2. Альтернативные попытки создания тезауруса: [PyТез](#);
3. Свой тезаурус конвертируют в форму, совместимую с WordNet;
4. Используется в Yandex, Rambler и т.д.;
5. Доступно по запросу <http://ruwordnet.ru/ru/>, [labinform.ru/...](http://labinform.ru/)

# PolarDB – новая платформа для хранения и обработки больших структурированных данных

Марчук Александр Гурьевич

1. NoSQL-База данных, разрабатывается открыто на [GitHub](#);
2. Высокопроизводительная, документ-ориентированная;
3. Необходимо описывать структуры хранимых документов (C#);
4. Реализация - .NET;
5. Использует Linq;
6. Коммерческое сопровождение не обеспечивают, не рекомендуют.

# Узорные структуры как математическая основа методов обнаружения знаний в предметных областях

Кузнецов Сергей Олегович

1. Развитие метода Анализа формального контекста (Formal Context Analysis);
2. Метод применен к анализу текстов из Интернета;
3. Разработан метод связывания предложений по местоимениям;
4. Результат представлен в виде древовидных структур;
5. Полученные деревья анализируются на предмет изъятия содержательной информации.

# Информационная система для принятия управленческих решений в сфере науки и образования

**Васенин Валерий Александрович**

1. Представлена информация о ИС "Истина" (МГУ);
2. Основная задача - наукометрия (внутренняя, для менеджмента научного персонала); Проблемы:
  1. Импорт и распознавание данных о публикациях (BibTeX, e-library, Scopus, WoS, Текст);
  2. Идентификация автора;
  3. Основной критерий - точность представления информации;
  4. Просят внедрить ...



# Компьютерная системная биология и биомедицина: текст майнинг, онтологии и базы знаний

Ак. Колчанов Николай Александрович и др.

1. Анализ текстов статей, опубликованных в рецензируемых журналах;
2. Восстановление текста из PDF;
3. Анализ текста, извлечение отношений;
4. Анализ отношений по заданным шаблонам;
5. Используется онтология в качестве модели для анализа;
6. Полученные результаты - знания (интерпретируемые данные);
7. Приложения полученных знаний для анализа влияния лекарств на больных диабетом.

# Теоретико-модельные методы извлечения и порождения знаний для построения онтологических моделей

Пальчунов Дмитрий Евгеньевич

1. Рассмотрена задача разработки онтологии рекомендательной системы для потребителей сотовой связи;
2. Использован метод Анализа формального контекста для построения системы знаний;
3. Система знаний представлена в виде иерархии по отношению к уровню общности понятия.

## Секционные доклады

# BIG DATA в поисковиках и лингво-частотный анализ языков по Улитке инноваций-путь моделирования смыслов.

Бадулин Николай Александрович

1. Инвестирует в проекты;
2. Фактически рассмотрел спиральную модель развития Стартапа и процесса руководства людьми;
3. Рассмотрены проблемы, возникающие на разных этапах;
4. К.Маркс, Ф.Энгельс, В.Ленин не были упомянуты.
5. Какое отношение к теме доклада ?

# Доказательное построение, верификация и синтез предикатных программ

Шелехов Владимир Иванович

1. Формализованное представление программ в предикативном языке;
2. Методология доказательства правильности программ базируется на АДТ, но АДТ требуемого качества не реализован;
3. Интересовался продвинутыми методами АДТ.

# Мультиагентный подход к разрешению кореференции при извлечении информации из текстов на основе онтологии

Сидорова Е. А., Гаранина Н. О., Серый А. С.

1. Основная задача - определить, какие понятия обозначают одно и то же в процессе анализа текстов;
2. Решение принимается на основе измерения схожести контекста и отношений, в которых состоят сравниваемые сущности;
3. Между понятиями анализируются отношения обобщения, эквивалентности, транзитивности и т.д.

# Новый алгоритм классификации потоковых данных

Анфилофьев Александр Евгеньевич

1. Решается задача классификации данных, приходящих в потоке (нельзя вернуться и снова что-то анализировать);
2. В процессе решения задачи система классификации дополнительно настраивается.

# Онтология научной графики в молекулярной спектроскопии

Привезенцев Алексей Иванович

1. Импорт данных в виде в т.ч. сканированных графиков;
2. Распознаются отдельные тренды (линии);
3. Вместе с Фазлиевым А.З. заинтересованы в сотрудничестве с нами.



# Опыт преподавания принципов методов и средств Semantic Web

Апанович Зинаида Владимировна

1. Рассмотрен магистерский курс по использованию формализованных онтологий для представления информации;
2. Рассмотрены вопросы использования Protege и др. сред проектирования онтологий;
3. Приведены примеры лабораторных работ: Интерпретация отношений, Построение запросов SPARQL.

# Использование онтологий при разработке продукционных экспертных систем

Дородных Никита Олегович, Юрин Александр Юрьевич

1. Рассмотрена задача создания графического веб-интерфейса для разработки экспертных систем;
2. В качестве методологии использования порождающего программирования использован подход, сходный с MDA;
3. На ранних этапах конструирования базы знаний используются онтологии.

# Использование технологий Linked Open Data при подготовке и публикации текстовых документов

к.т.н. Черкашин Е.А.

лаб. 4.1. Комплексных информационных систем

ИДСТУ СО РАН

к.т.н. Шигаров А.О., к.ф.-м.н. Орлова И.В.,

к.б.н. Михайлов И.С.

ИДСТУ СО РАН, ИРННТУ,

ЛИН СО РАН, ИНЦ СО РАН

# Цели и задачи

Средствами LOD + HTML5 в документе формируется логическая разметка, HTML5 + CSS + алгоритмы - различные варианты визуального представления.

Цель - разработка ПО для подготовки и верстки документов, размеченных согласно LOD и порождаемых при помощи LOD,

- позволяющее создавать документы на рабочей станции пользователя, загружая и преобразуя фрагменты текста из других документов;
- реализующее различные представления (View): форма ввода и др.
- Интеграция с другими приложениями обработки LOD.

# Рабочая программа дисциплины



МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ  
ФЕДЕРАЦИИ  
федеральное государственное бюджетное образовательное учреждение  
высшего образования  
«ИРКУТСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»  
ФГБОУ ВО «ИГУ»  
Институт математики экономики и информатики

Кафедра информационных технологий

УТВЕРЖДАЮ

Директор ИМЭИ

" " 20 г.

Рабочая программа дисциплины (модуля)  
Б1.В.ДВ.3.1. Технологии программирования

Направление подготовки:	10.03.01 (090900) Информационная безопасность
Направленность (профиль)	- общий
Квалификация (степень) выпускника	- бакалавр
Форма обучения	- очная

Иркутск 2016 г.

Согласовано с УМК факультета (института)

Рекомендовано кафедрой:

Протокол № от " " 20 г.

Протокол № от " " 20 г.

Председатель (подпись)

Зав. кафедрой (подпись)

Содержание

1. Цели и задачи дисциплины (модуля)
2. Место дисциплины в структуре ОПОП
3. Требования к результатам освоения дисциплины (модуля)
4. Объем дисциплины (модуля) и виды учебной работы (разделяется по формам обучения)
5. Содержание дисциплины (модуля)
6. Перечень семинарских, практических занятий и лабораторных работ
7. Примерная тематика курсовых работ (проектов)
8. Учебно-методическое и информационное обеспечение дисциплины (модуля)
9. Материально-техническое обеспечение дисциплины (модуля)
10. Образовательные технологии
11. Оценочные средства (ОС)

## 1. Цели и задачи дисциплины (модуля)

Целью преподавания дисциплины «Технологии программирования» является освоение студентами практических навыков в области разработки программного обеспечения на основе современных подходов к проектированию сложных, гетерогенных, распределенных информационных систем. Развитие навыков системного мышления, необходимого для

# Редактор Dokie.ly

## C.Saradisly,..., T.-B. Lee, 2017

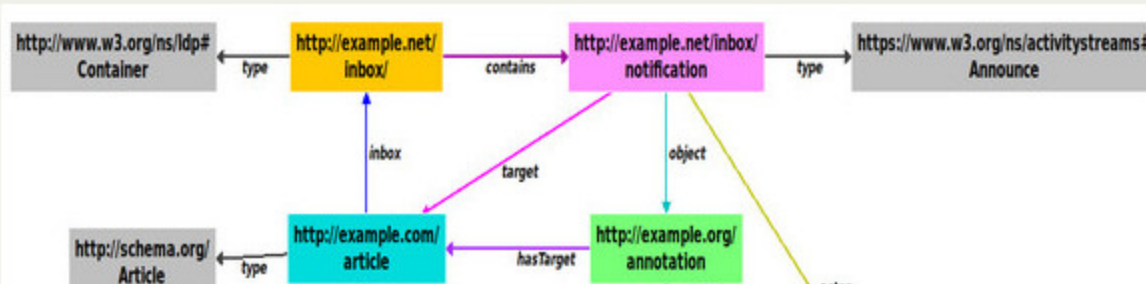
dokieli as a client-side application can be deployed on a single-page or through a browser extension, which can consume and interact with Linked Data (or otherwise) anywhere on the Web. We consider an HTML document with CSS and JavaScript as the default UI of a document. Through the separation of the layers, other UIs can work in conjunction with the same HTML. This approach is independent from specific server-side software, proprietary APIs or the requirement to have an account.

On the other hand, if desired and available, users can participate using their own profiles (WebIDs) located anywhere on the Web, and get to store and make their annotations in their own personal storage, as well as assign access controls to documents. Similarly, a decentralised communications protocol, Linked Data Notifications (LDN) – W3C Proposed Recommendation – is used to get past the limits of centralisation by enabling communication to happen across independent servers and applications. Figure 2 depicts relations between the kinds of entities which underlay dokieli's architecture, where nodes are under different domains and authority.

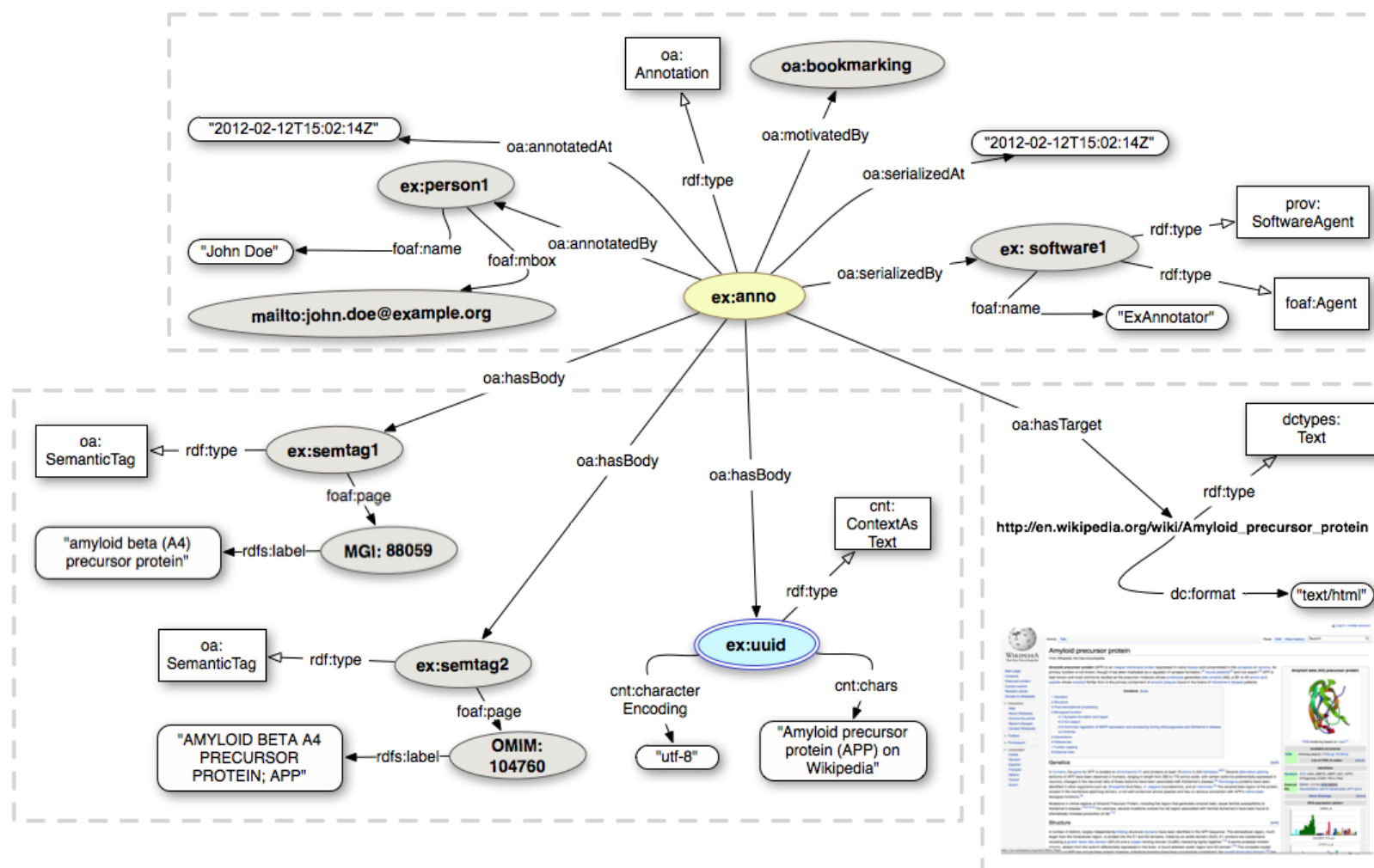
Samurai Cat

2017-04-28 16:33:12

Sadly the ideal (the fantasy) is undermined by the reality. '... if desired and available ...' think for a moment what that means. Desired would mean that someone has made a decision to 'participate' knowing the process. Available would mean that this participation is possible, using, as you suggest, WebID. I think that as it stands WebID is just too complex to be either wanted by users (it cannot be understood) or available (there are no implementations that I can find that suggest, without much confusing tinkering, a plausible PoC).



# Сандарт Open Annotation



# Аннотация документа

. . <body> . . . .

<main lang="ru" resource="#annotation"

typeof="oa:Annotation"

id="main-document-container">

<div property="oa:hasTarget"

resource="#course-work-program"></div>

<article property="oa:hasBody"

typeof="schema:Article foaf:Document

curr:WorkingProgram"

resource="#course-work-program" id="main-document">



# Данные курса

## Учебный план специальности 01.03.02 Прикладная математика и информатика

### 1. Общие сведения учебного плана

#### Сведения по Учебному плану

Профиль подготовки: Математическое и компьютерное моделирование в технике и экономике, методы принятия решений

#### Сведения о кафедре, разработавшей Учебный план

Кафедра: Математического анализа и дифференциальных уравнений,  
Факультет: ИМЭИ.

#### Сведения о специальности

Квалификация: Бакалавр

Форма обучения: очная

Программа подготовки: прикладн. бакалавриат

#### Руководители

Проректор по учебной работе: Не распознан

Начальник УМУ: А.И. Вокин

Директор: М.В. Фалалеев

### 2. Список компетенций

#### Дисциплина: Б1.В.ДВ.3.1. Технологии программирования

- способность приобретать новые научные и профессиональные знания, используя современные образовательные и информационные технологии (ОПК-2)
- способность критически переосмысливать накопленный опыт, изменять при необходимости вид и характер своей профессиональной деятельности (ПК-3)
- способность к разработке и применению алгоритмических и программных решений в области системного и прикладного программного обеспечения (ПК-7)

### 3. Список курсов специальности

- Б1.Б.3 «Философия»

# Верстка

- методиками экстремального и *agile*-программирования.

26

## 4. Объем дисциплины (модуля) и виды учебной работы (разделяет формам обучения)

Вид учебной работы	Всего часов / зачетных единиц	Семестры	
		3	4
Аудиторные занятия (всего)	108	33	75
в том числе:			
Лекции	36		36
Практические занятия (ПЗ)			
Семинары (С)			
Лабораторные работы (ЛР)	66	30	36
КСР	6	3	3
Самостоятельная работа (всего)	45	39	6
в том числе:			
Курсовой проект (работа)			
Расчетно-графические работы			
Реферат (при наличии)			
Другие виды самостоятельной работы			
Изучение учебной, научной и методической литературы с привлечением электронных			

# Использованные онтологии

**Friend-of-a-friend (foaf)** - информацию об агентах: физических, юридических, программных.

**Provenance (prov)** - происхождение информации: связь между документами.

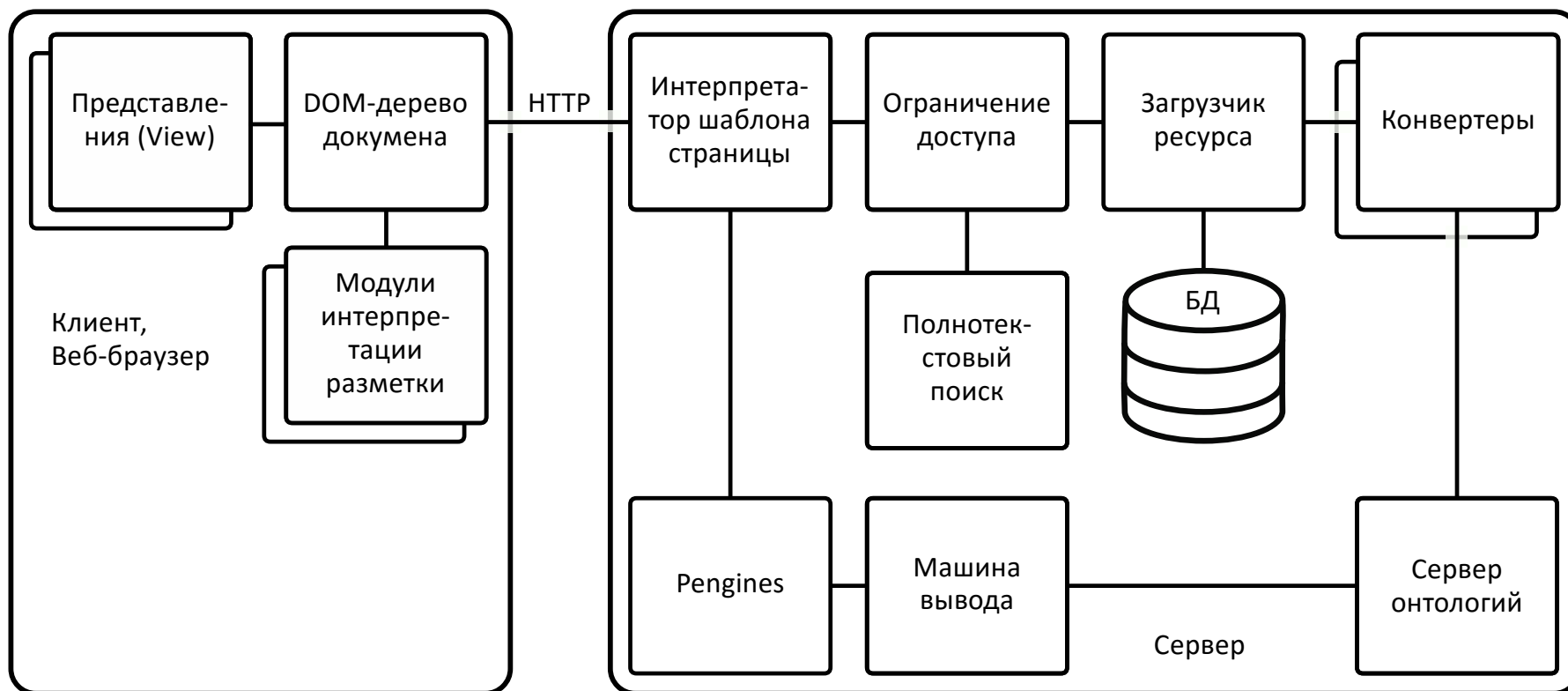
**Dublin Core (dc)** - разметка редактируемой аннотации.

**DBPedia resource (dbr)** – ссылки на экземпляры объектов и классы.

**Schema.org (schema)** - объекты, распознаваемые поисковыми агентами Google, Yandex, Yahoo и др. (часть-целое).

**The Bibliographic Ontology (bibo)** - разметка списка литературы.

# Архитектура системы



# Области применения и дальнейшее развитие

- Автоматизация подготовки документов;
  - Редактор, учитывающий контекст;
  - Экспорт в офисные форматы;
  - Самоорганизующийся документооборот;
- Исходные данные для электронных курсов;
  - Импорт в системы автоматизации обучения;
- Автоматизация разметки
  - Входящей документации (ФОС);
  - Существующих учебных программ.

# Вопросы

<https://github.com/isu-enterprise/isu.aquarium>