

Федеральное агентство научных организаций  
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ УЧРЕЖДЕНИЕ  
НАУКИ ИНСТИТУТ ПРОБЛЕМ РЫНКА  
РОССИЙСКОЙ АКАДЕМИИ НАУК  
(ИПР РАН)

УДК 001.89 : 004.774.2

№ госрегистрации

АААА-А18-118011990112-8

Инв. №

"УТВЕРЖДАЮ"

Директор ИПР РАН

член- корр. РАН

В.А. Цветков

21 » декабря 2017 г.

ОТЧЕТ  
О НАУЧНО-ИССЛЕДОВАТЕЛЬСКОЙ РАБОТЕ

Сущность и выразительные средства семантического  
аннотирования

по теме:

ИССЛЕДОВАНИЕ, РАЗРАБОТКА И ПОДДЕРЖКА  
ВЕБ-ИНФРАСТРУКТУРЫ НАУЧНЫХ ИНФОРМАЦИОННЫХ РЕСУРСОВ  
С ОТКРЫТЫМ ДОСТУПОМ

№ 0163-2016-0006 Государственного задания ИПР РАН

(промежуточный)

Руководитель темы  
зав. лабораторией,  
к.т.н., доцент

*М.Р. Когаловский*  
2017.12.12

М.Р. Когаловский

МОСКВА – 2017

## СПИСОК ИСПОЛНИТЕЛЕЙ

Руководитель темы,  
исполнитель,  
зав. лабораторией,  
к.т.н., доцент

*М.Р. Коголовский*  
*2017.12.12*

М.Р. Коголовский (все разделы)

Исполнитель, н.с.

*Е.М. Ильменская*  
*2017.12.12*

Е.М. Ильменская (раздел 5)

## РЕФЕРАТ

Отчет 27 с., 5 табл., 16 источников

ИНФОРМАЦИОННЫЙ РЕСУРС, АННОТАЦИЯ, ОБЩАЯ МОДЕЛЬ АННОТАЦИИ, СЕМАНТИЧЕСКАЯ АННОТАЦИЯ, ТАКСОНОМИЯ, ЦИТИРОВАНИЕ, ЭЛЕКТРОННАЯ БИБЛИОТЕКА, СИСТЕМА СОЦИОНЕТ

В настоящее время активно развивается тенденция повышения семантического уровня представления и оперирования информационными ресурсами в информационных системах, прежде всего, применяемых в научных исследованиях. Благодаря этому существенным образом повышается эффективность использования таких ресурсов. Это приобретает особую актуальность в связи с доминантой четвертой парадигмы в научных исследованиях – явления, при котором интенсивное использование данных становится основой научных исследований, и родилось новое направление в информатике, названное Data Science. Повышение семантического уровня представления информационных ресурсов может быть достигнуто различными способами. Один из подходов состоит в использовании их семантического их аннотирования, т.е. ассоциирования с информационными объектами явным образом представленных данных, описывающих их семантику. По указанной причине исследование существа и инструментария семантического аннотирования является актуальной проблемой.

Семантическое аннотирование существенным образом обогащает контент научной публикации, может иметь различный характер и может служить различным целям. В частности, его можно использовать для обеспечения семантического поиска аннотаций и аннотируемых фрагментов нужной категории в тексте публикаций, для оценки содержания отдельных фрагментов текста, для обогащения их контента путем различного рода дополнений к ним, а также для их классификации. В электронных библиотеках семантическое аннотирование библиографических ссылок в текстах публикаций на используемые источники позволяет генерировать новые наукометрические показатели.

В отчете рассматривается существо семантического аннотирования, определяются основные понятия, исследуется общая модель семантической аннотации, обсуждается таксономический подход к представлению семантики аннотаций, приводятся примеры таксономий, которые основаны на различных свойствах аннотаций и которые могут использоваться для этих целей. В качестве примера кратко рассматривается реализация семантического аннотирования в научной информационной системе Соционет.