

Онтологическое аннотирование библиографических ссылок в научных публикациях

М.Р. Когаловский
Институт проблем рынка РАН, Москва, Россия

Ключевые слова: библиографическая ссылка, онтологическое аннотирование, наукометрия

В работе предлагается подход, предусматривающий обогащение определяемой явным образом семантикой традиционных «немых» ссылок цитирования в научных публикациях и тем самым создание новых более информативных источников данных для наукометрических исследований.

Актуальность измерений различных видов индексов цитирования научных публикаций в последние годы существенно возросла. Индексы цитирования стали ключевым критерием оценки вклада ученого в развитие науки. На основе статистики цитирования измеряются и импакт-факторы периодических изданий, которые рассматриваются как характеристики их научного авторитета. Однако такие оценки обладают существенным недостатком. Действительно, могут возникать парадоксальные ситуации, когда высоким индексом цитирования (а, значит, в соответствии со сложившейся практикой, и высокой степенью значимости) обладает статья, содержащая грубые ошибки и/или принципиальные заблуждения и в связи с этим вызывающая активный отклик научного сообщества. Эта проблема возникает в связи с тем, что сложившиеся методы наукометрии опираются на библиографические ссылки, которые сами по себе *не несут какой-либо информации*, характеризующей цель цитирования, мотивы, побудившие автора цитирующей работы использовать такие источники, его мнение о них и т.п. Поэтому такие ссылки мы называем «*немыми*». Они характеризуют лишь факт существования *связи цитирования* между рассматриваемыми публикациями.

Использование *семантики ссылок* в наукометрических измерениях позволяет избежать указанных выше ситуаций благодаря возможности формирования статистики цитирования, учитывающей мотивы цитирования публикаций и тем самым реальную степень их значимости в развитии научных знаний. Для спецификации семантики библиографических ссылок в данной работе предлагается подход, при котором эксперт анализирует текст публикации и осуществляет аннотирование встречающихся в нем ссылок на используемые источники (*внутритекстовых ссылок, знаков сноски, выноски и отсылки*) на основе контекстной информации и *онтологии семантических связей*, которая определяет представляющие интерес виды отношений цитирования между

научными публикациями. При компьютерной обработке текста публикации с аннотированными ссылками из него будут извлекаться данные о связях данной публикации с цитируемыми источниками, в том числе и о семантике этих связей. Эти данные могут коллекционироваться наряду с аналогичными данными о других публикациях, полученными иным образом, и служить источником для наукометрических исследований, более содержательных по сравнению с традиционными.

Важно отметить, что в публикации может использоваться *множество ссылок на один и тот же источник, обладающих различной семантикой*. Все эти связи учитываются при данном подходе в отличие от традиционных индексов цитирования, учитывающих только одну «немую» связь между двумя публикациями, выявленную при анализе приставейных списков использованных источников.

В данной работе обсуждаются основные принципы предлагаемого подхода к онтологическому аннотированию связей цитирования и его реализации. Приводится краткая характеристика существа семантического и, в частности, онтологического аннотирования. Обсуждаются принятые в издательской деятельности способы представления ссылок на используемые источники в научных публикациях, уточняются объекты аннотирования и рассматриваются средства спецификации их семантики, характеризуется среда реализации предлагаемого подхода – система Соционет.

Отмечается целесообразность *стандартизации* онтологии библиографических ссылок и форматов представления их онтологических аннотаций.

Литература

1. ГОСТ 7.0.5-2008 Система стандартов по информации, библиотечному и издательскому делу. Библиографическая ссылка. Общие требования и правила составления. Издание официальное. М.: Стандартинформ, 2009.

2. Когаловский М.Р. Метаданные в компьютерных системах // Программирование. МАИК/Наука «Интерпериодика». 2013. Т. 39, № 4. С. 28-46.

3. Когаловский М.Р., Паринов С.И. Классификация и использование семантических связей между информационными объектами в научных электронных библиотеках // Информатика и ее применения. 2012. Т. 6. Вып. 3. С. 32-42.

4. Когаловский М.Р., Паринов С.И. Семантическое структурирование контента научных электронных библиотек на основе онтологий. В кн.:

"Современные технологии интеграции информационных ресурсов: сборник научных трудов». Вып. 2. Санкт-Петербург: Президентская библиотека им. Б.Н. Ельцина, 2011. С. 26-45.

5. Онлайн-научная инфраструктура Соционет. <http://socionet.ru/>

6. Паринов С.И., Когаловский М.Р. Технология семантического структурирования контента научных электронных библиотек. Труды XIII Всероссийской научной конференции «Электронные библиотеки: перспективные методы и технологии, электронные коллекции – RCDL-2011. Воронеж, 19-22 октября 2011 г.». г. Воронеж: Воронежский государственный университет, 2011. С. 94-103.