



Семантическая сеть и семантическая паутина

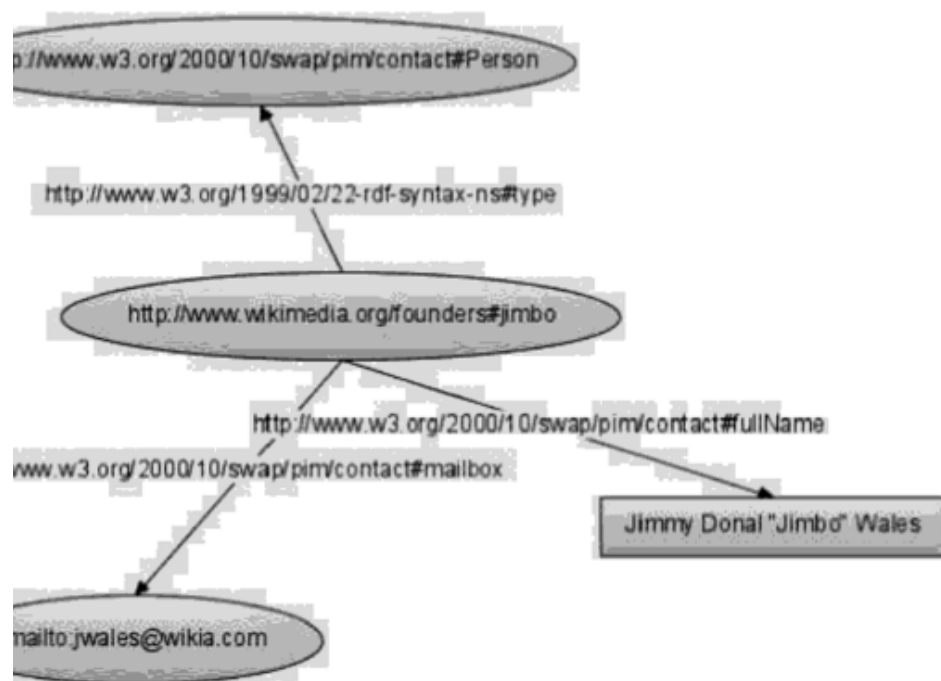
Как показано в подразд.7.2, семантическая сеть имеет давнюю историю, но последнее время это понятие не вызывало никаких мысленностей. Однако, после того, как изобретатель Всемирной паутины (WWW) Тим Бернерс Ли в 2001 году провозгласил концепцию “Semantic Web”, в русскоязычной литературе эти два понятия стали сливаться. Для устранения недоразумений предлагается переводить понятие “Semantic Web” как «Семантическая Паутина» или, в данном случае — просто «Паутина». Суть идеи Тима Бернерса Ли заключается в том, чтобы снабдить ресурсы Интернета специальными метаданными, пригодными для компьютерной обработки и однозначно характеризующими свойства и содержание ресурсов Всемирной паутины, вместо используемого в настоящее время текстового анализа документов. В частности, предполагается снабдить все ресурсы Интернета тегами, позволяющими установить автора каждого документа. Стоит отметить, что это решение сделано «вдогонку» к концепции WWW после того, как Интернет из хранилища знаний превратился в информационную «помойку» из-за допускаемой анонимности ресурсов. Именно данное свойство семантической паутины нас интересует в первую очередь. В центре нашего интереса — свойства семантической Паутины доставлять пользователю информацию, которую он хочет получить.

Основное отличие Семантической Паутины от WWW заключается в том, что связи между ресурсами WWW отражают размещение документов, а связи Паутины — содержание. Правда, как в WWW, так и в Паутине ссылки содержат адреса ресурсов, но в WWW ссылки анонимные; их смысл должен понимать пользователь. В Паутине же ссылки имеют явно указанный смысл, который доступен для машинной обработки. Таким образом, Семантическая Паутина подразумевает организованный поиск информации вместо имеющего место сейчас

серфинга, когда пользователь Интернета переходит по ссылкам от мента к документу.

Семантическая паутина: принципы и текущее состояние

Основой Семантической Паутины лежит тот же триплет субъект — предикат — объект. Отличие заключается в представлении каждого из элементов триплета. В Паутине субъект, объект и предикат представлены универсальными идентификаторами ресурсов или URI (Uniform Resource Identifier). Так, граф визитной карточки основателя Википедии выглядит следующим образом (рисунок позаимствован в той же Википедии):



И только выглядит как адрес ресурса URL. На самом деле по любому приведенных на рисунке адресов, таких как `http://www.w3.org/2000/10/swap/pim/contact#mailbox`, вы не найдете ничего похожего на почтовый ящик, или хотя бы его описание, как это может показаться на первый взгляд. Такая система именования нужна только для обеспечения уникальности имен.

Более популярным языком представления знаний в Паутине является язык RDR Фрагмент RDF-документа, описывающие отношения. Например, если Сергей имеет ребенка Никиту, а Никита имеет родителя Сергея, то это можно записать так:

Этот формат не предназначен для чтения человеком. Редакторы семантических документов, например, наиболее распространенный редактор онтологий Protege (<http://protege.stanford.edu/>), дают

ожность вводить и редактировать факты в экранных формах
ймах), как показано ниже.

настоящее время создание RDF/OWL документов ведется
льными энтузиастами. Количество документов в Интернете
лико, и их можно найти с помощью специального поискового
эра, например, SWOOGLE (<http://swoogle.umbc.edu/>). Развитие
кта сдерживается по многим причинам. В основном это отсутствие
рой и прямой выгоды от создания семантических ресурсов,
ность формализации знаний и отсутствие универсальных агентов
извлечения знаний. Представленные в Интернете реализации
плектуальных агентов ограничиваются студенческими разработками
е путеводителя по пивным города Саутхемптон ([http://
:twine.eom/item/llby40gxx-pl/southampton-pub-guide-in- rdf](http://twine.eom/item/llby40gxx-pl/southampton-pub-guide-in-rdf)) и
ого агента, советующего, какое вино лучше употребить с каждым из
 ([http://ksl.stanford.edu/people/dlm/ webont/wineAgent/](http://ksl.stanford.edu/people/dlm/webont/wineAgent/)).