

## Из истории Семантического Веб

*Верхова Ольга Геннадьевна*

*Брянский государственный университет имени академика Ивана Георгиевича Петровского*  
*Магистрант*

*Иванова Наталья Александровна*

*Брянский государственный университет имени академика Ивана Георгиевича Петровского*  
*Кандидат технических наук, доцент кафедры информатики и прикладной математики*

### **Аннотация**

Данная статья посвящена обзору исторического развития Семантического Веб. Проведенное исследование демонстрирует развитие языков разметки, и открывает возможные перспективы данного направления Web 3.0.

**Ключевые слова:** Семантическая паутина, Семантический Веб, Тим Бернерс-Ли, RDF.

## The history of the Semantic Web

*Verkhova Olga Gennadyevna*

*Bryansk State University named after Ivan G. Petrovsky*  
*Master*

*Ivanova Natalia Aleksandrovna*

*Bryansk State University named after Ivan G. Petrovsky*  
*Ph.D., assistant professor of computer science and applied mathematics*

### **Abstract**

This article provides an overview of the historical development of the Semantic Web. The study shows the development of markup languages, opens up opportunities and prospects of this direction Web 3.0.

**Keywords:** Semantic Web, Tim Berners-Lee, RDF.

Чаще всего при поиске информации в сети необходим конкретный ответ на вопрос, а не список документов, которые, возможно, его содержат. Для того чтобы компьютеры «научились» понимать смысл документов и данных, была предложена Семантическая паутина (Semantic Web, Семантический Веб).

Первой предпосылкой к возникновению концепции Semantic Web стало создание в середине 80х годов XX века программы для управления

документами Enquire. Ее создателем стал Тим Бернерс-Ли, сотрудник ЦЕРН (Европейской организации по ядерным исследованиям), ставший создателем Всемирной паутины, а в данный момент возглавляющий Консорциум Всемирной паутины. Enquire позволяла хранить информацию о данных и связях между ними [1].

Важным этапом, предшествующим появлению Семантического Веба, является создание World Wide Web (WWW) – Всемирной паутины. В 1989 году Тим Бернерс-Ли опубликовал документ «Информационный менеджмент: Планы», определивший возможность создания Всемирной паутины.[2] Закономерным результатом манифеста стало появление в 1990 году WWW-сервера и первого браузера.

В 1994 году возникла международная организация – Консорциум Всемирной паутины W3C (WorldWideWebConsortium), – главная цель которой разрабатывать и внедрять стандарты для WWW [3].

В 1995 году в городе Дублин, штат Огайо, прошел семинар NSCA (Национальный центр суперкомпьютерных приложений), на котором был предложен набор из пятнадцати элементов метаданных, описывающих информацию на веб-ресурсах. Данный набор получил название «Дублинское ядро», и в дальнейшем получил широкое распространение, в том числе и в государственных стандартах [4].

В 1997 году началась разработка технологии Rich Site Summary RSS – формата для описания различных изменений на веб-ресурсах. Впоследствии версии технологии 0.9 и 1.0 основывались на формате Resource Description Framework RDF.

В сентябре 1998 года Тим Бернерс-Ли опубликовал статью «Дорожная карта Семантического Веб», представляющую собой особый путь дальнейшего развития технологий Всемирной паутины [5].

В 1998 году была разработана спецификация Extensible Markup Language 1.0 (XML), представляющая метаязык, который лег в основу разработки других языков разметки. XML создавался как инструмент для удобного чтения и обработки документов как компьютером, так и человеком, и был ориентирован на Интернет. Так же была предложена Resource Description Framework Schema (RDF Schema), описывающая классы и свойства для разработанной впоследствии модели представления данных.

В 1999 году Консорциум Всемирной паутины создал специализированные группы по проектированию Семантического Веба. Тогда же им была опубликована первая версия спецификации Resource Description Framework (RDF), впоследствии дважды переработанная в 2004 и 2014 годах. RDF представляет собой модель для обмена данными по сети. Он разработан так, чтобы при развитии систем не было необходимости изменять принципы обработки данных. [6].

В 2000 году были начаты исследования по описанию онтологий (описания классов объектов, их свойств и взаимоотношений в определенной предметной области). Тогда был создан семантический язык разметки DAML+OIL, основывающийся на технологиях RDF и RDF Schema. В нем

применяются расширенные примитивы моделирования. В 2001 году Тим Бернерс-Ли в журнале «Сайентифик Америкэн» опубликовал статью, раскрывающую возможности Семантического Веб обществу [7, 8].

В 2004 году спецификация RDF была переработана в шесть отдельных спецификаций:

- 1) RDF Primer (содержит начальные сведения о RDF);
- 2) RDF Concepts and Abstract Syntax (описывает понятия и абстрактный синтаксис);
- 3) RDF/XML Syntax Specification (описывает нормативный синтаксис сериализации графа RDF как документа XML);
- 4) RDF Semantics (описание семантики и полной системы правил для RDF и RDF Schema);
- 5) RDF Vocabulary Description Language 1.0: RDF Schema (описывает встроенные в RDF словари, а также как создавать новые словари);
- 6) RDF Test Cases (содержит набор тестовых данных).

Также в этом же году был представлен язык OWL – язык описания онтологий, разработанный специально для Семантической паутины, и введено понятие Web 2.0, описывающее Интернет как систему, привлекающую пользователей для наполнения ее данными.

В 2005 году появился проект Friend of a Friend (FOAF), или «Друг друга» – экспериментальная информационная система, основанная на одноименном языке FOAF. FOAF основан на RDF, а его определения на OWL, и состоит из терминов, описывающих различные критерии, применимые к людям, например имя, интересы, адрес электронной почты [9].

В 2006 году Тим Бернерс-Ли сделал заявление, согласно которому Семантический Веб должен стать непосредственной частью Web 3.0, нового этапа развития Всемирной паутины, когда не только пользователи смогут самостоятельно наполнять Интернет содержимым, но и произойдет такое развитие технологической и социокультурной платформ, которое позволит создавать качественный профессиональный контент с отдельным автоматическим обменом данными между сайтами. Также появился язык запросов SPARQL (SPARQL Protocol and RDF Query Language).

В это же время стартовал проект GeoNames – географическая база данных, основанная на RDF. По состоянию на 2015 год она насчитывает более 10 миллионов наименований [10].

В 2007 году был запущен проект DBpedia, также основанный на стандарте RDF. Его суть заключается в извлечении структурированной информации из данных открытой энциклопедии Википедии, а затем опубликования результатов под свободной лицензией. Сейчас DBpedia описывает более 4,58 миллионов понятий [11].

В октябре 2007 года стартовал проект «Облако связанных открытых данных» (Linked Open Data Cloud). Проект объединяет различные системы, использующие RDF, и визуализирует используемые ими взаимосвязанные наборы данных в виде графа. На рисунке 1 показана визуализация за 2007 г. [12].

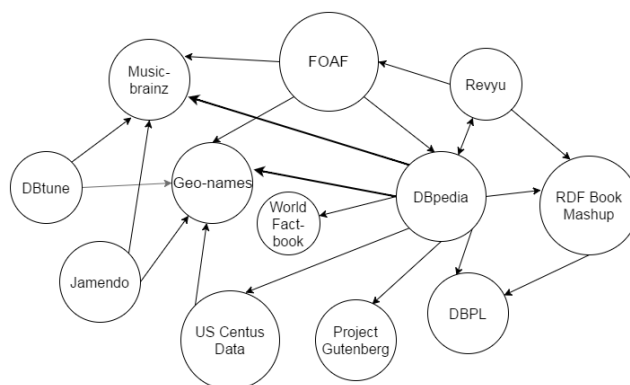


Рисунок 1 – Облако связанных открытых данных за 2007 г.

В 2008 году версия SPARQL 1.0 стала стандартом. Она включала в себя непосредственно язык запросов, протокол, а также формат выдачи результатов XML. В 2009 году версия языка OWL2 становится рекомендацией Консорциума Всемирной паутины [10].

В 2013 году версия SPARQL стала рекомендацией Консорциума Всемирной паутины.

В 2014 году спецификация RDF была обновлена: введены понятия наборов данных и графов RDF, дополнены некоторые аспекты языка.

К 2015 году проект Linked Open Data Cloud достиг значительных масштабов. На рисунке 2 представлена визуализация взаимосвязей систем по состоянию на 2015 г [13]. По сравнению с 2007 годом облако выросло более чем в 47 раз.

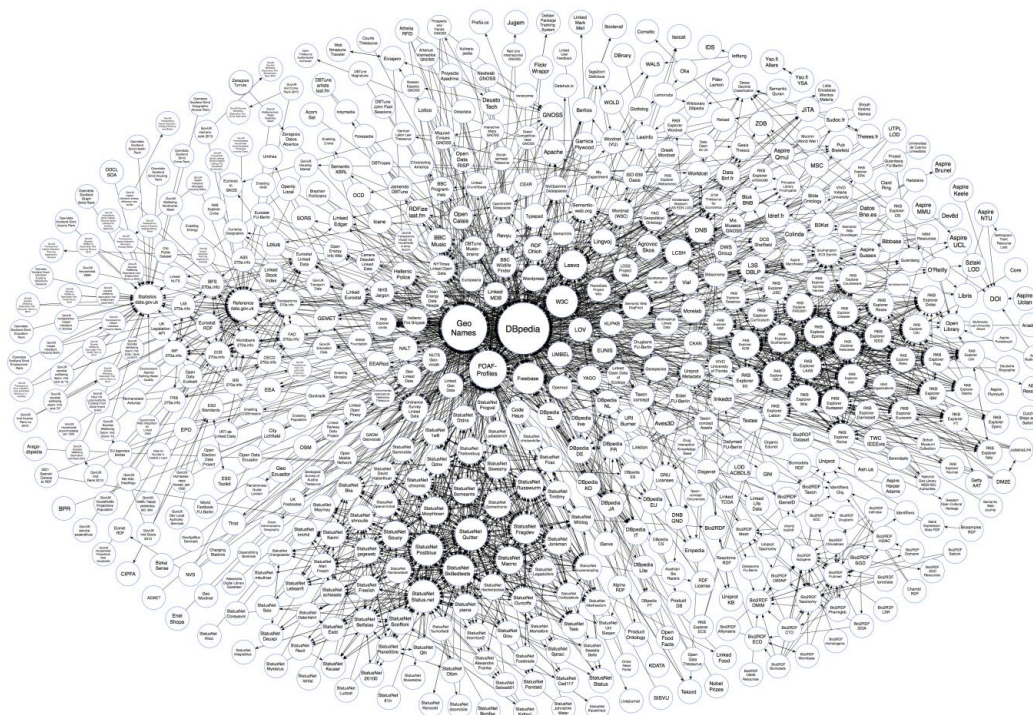


Рисунок 2 – Облако связанных открытых данных за 2015 г.

В настоящий момент Семантическая паутина постепенно развивается, и все больше исследователей считают, что переход к Семантическому веб необходим и неизбежен.

Уже сформирован базис технологий, позволяющих его использовать, постепенно развиваются семантически ориентированные ресурсы. Можно ожидать, что с дальнейшим расширением Семантической паутины удастся стереть границы между отдельными ресурсами и создать единое информационное пространство связанных данных.

### Библиографический список

1. Тим Бернерс-Ли [Электронный ресурс] – URL: <http://chernykh.net/content/view/510/706/> (Дата обращения 26.02.2016).
2. Первоначальный план WWW [Электронный ресурс] – URL: <https://www.w3.org/History/1989/proposal.html> (Дата обращения 26.02.2016).
3. Ахиллесова пята Семантического Веба [Электронный ресурс] – URL: <http://old.computerra.ru/magazine/362912/> (Дата обращения 27.02.2016).
4. Федотов А.М., Барахнин В. Б. Проблемы поиска информации: истории и технологии [Электронный ресурс] – URL: [http://www-sbras.nsc.ru/win/elbib/data/dinamo\\_cat/1.pdf](http://www-sbras.nsc.ru/win/elbib/data/dinamo_cat/1.pdf) (Дата обращения 20.02.2016).
5. Semantic Web Road map [Электронный ресурс] – URL: <http://www.w3.org/DesignIssues/Semantic.html> (Дата обращения 26.02.2016).
6. RDF [Электронный ресурс] – URL: <http://www.w3.org/RDF/> (Дата обращения 26.02.2016).
7. Семантический Веб [Электронный ресурс] – URL: <http://yury.name/internet/08ia.pdf> (Дата обращения 26.02.2016).
8. DAML+OIL (March 2001) Reference Description [Электронный ресурс] – URL: <http://www.w3.org/TR/2001/NOTE-daml+oil-reference-20011218> (Дата обращения 27.02.2016).
9. FOAF [Электронный ресурс] – URL: <http://www.foaf-project.org/> (Дата обращения 27.02.2016).
10. SPARQL 1.1 Overview [Электронный ресурс] – URL: <http://www.w3.org/TR/sparql11-overview/> (Дата обращения 26.02.2016).
11. DBpedia [Электронный ресурс] – URL: <http://wiki.dbpedia.org/> (Дата обращения 28.02.2016).
12. Проект Linked Open Data Cloud [Электронный ресурс] – URL: <http://lod-cloud.net/> (Дата обращения 25.02.2016).
13. Linking Open Data [Электронный ресурс] – URL: <http://www.w3.org/wiki/SweoIG/TaskForces/CommunityProjects/LinkingOpenData> (Дата обращения 24.02.2016).