

Ефимов А.А., г. Екатеринбург

Семантическая паутина - Интернет будущего

Заводя речь о концепции семантической паутины (Semantic Web), стоит сначала определить некоторые термины, разобраться в том, что же такое – семантическая паутина.

Говоря «Интернет», обычно, мы понимаем под этим словом некую застывшую в развитии технологическую единицу, тем самым сильно упрощая истинное положение дел. На самом деле — что, впрочем, тоже не секрет — «сеть» прошла два этапа своей эволюции, сделавших её такой, какой мы видим её сегодня. Первым этапом, уходящим корнями в 70-ые годы, был собственно Интернет — выросшая из военного проекта глобальная компьютерная сеть, узлы которой обменивались данными, не

требовавшими обработки в реальном времени (письма, файлы и т.п.). В начале 90-ых родилась World Wide Web — Всемирная паутина, само название которой подразумевает упрощение и популяризацию технологий, бывших уделом учёных, военных и компьютерных гениев. Web предоставил унифицированный способ доступа к различным типам данных. В результате этого, вместилищем информации стали не файловые архивы, а web-сайты, работа online сделалась проще, наглядней и интересней, а сама «сеть» превратилась из глобальной компьютерной в глобальную информационную. Но и Web, по сути, без особых изменений доживший до настоящего времени, не стал последним этапом: на очереди следующая эволюционная ступень, именуемая Semantic Web — сеть семантическая [<http://www.w3.org/2001/sw>].

Семантическая паутина была задумана консорциумом W3 [<http://www.w3.org/>] очень давно. С середины 90-х публиковались разные статьи и заметки, которые не привлекали особого внимания широкой общественности, но однажды, председателем консорциума W3 была опубликована статья, обратившая на себя внимание всего Интернет сообщества:

17 мая 2001 года в журнале Scientific American появилась статья Тима Бернерса-Ли (Tim Berners-Lee), Джеймса Хендлера (James Handler) и Оры Лассила (Ora Lassila) под названием «The Semantic WEB»¹.

Надо отметить, что Тим Бернерс-Ли — это отец WWW. Он не только первым предложил концепцию всемирной паутины, но и воплотил её в жизнь. К словам такого человека просто нельзя не прислушиваться.

Стиль описания напомнил мне громкие пророчества манифеста «Some Computer Science Issues in Ubiquitous Computing» [3], которые, кстати, сбываются, про то, что главной проблемой является не повсеместная компьютеризация жизни, а конфиденциальность личной жизни в таких условиях. Или, опять же, сюжеты про цифровые дома из японской анимационной фантастики — во время того, как ты сидишь на работе, электронная штучковина внутри твоего скромного жилища находит оставленное

¹ Доступен перевод [http://ezolin.pisem.net/logic/semantic_web_rus.html], выполненный Евгением Золиным.

задание «приготовить компот», холодильник через сеть связывается с ближайшим гастрономом, вместе они решают, что для приготовления необходимы сухофрукты и вода, после этого прямо к холодильнику, с ближайшего рынка, т.к. это ровно на 35.987% выгоднее гастронома, доставляются отборные сухофрукты. Кухонный процессор снова идет в сеть, ищет рецепт компота из сухофруктов, и его не ставит в тупик то обстоятельство, что рецептов несколько, он то знает не только предпочтения хозяина, но и его возраст, состояние здоровья и многое другое, поэтому находит-таки рецепт компота разжижающего кровь «для тех, кому до шестнадцати и старше». Цифровой дом так же находит в сети информацию о том, что для подачи этого напитка стол должен быть сервирован стаканом, причем лучше высоким, с прямыми стенками...

Впечатляющий футуризм. У этой статьи, как и у всех подобных была только одна цель — привлечь внимание к Semantic Web всех, кого только можно было. Зачем — непонятно. Скорей всего, авторы просто сделали открытие (а Semantic Web на самом деле впечатляющее открытие), и, находясь, некоторое время в просветленном состоянии решили написать все, что думают.

В двух словах, под семантической паутиной понимается структура, наполненная информацией о смысловом содержании, понятной машинам. В самом деле, задумайтесь: ведь современные web-страницы созданы для человека, компьютерным программам не под силу проанализировать их содержимое — искусственный интеллект ведь пока не изобрели. Главной задачей семантической паутины станет добавление специальных данных, которые позволят компьютерным программам понимать web-страницы. Скромное на первый взгляд добавление выведет Интернет на новый виток эволюции. Представьте, что браузеры научатся «понимать», о чём говорится на web-страничках, которые они открывают. Это сразу же на порядок повысит эффективность работы с web-ресурсами: компьютеры сами, без помощи человека смогут находить нужную нам информацию, самостоятельно отсеивать сор, похожий словарно, но отличный по смыслу, будут в состоянии автоматически выполнять самые различные поручения по работе с web-контентом, вроде заказа мест в ресторане или подбора рефератов. Человеческий интеллект, который сегодня является необходимым инструментом

при работе с Сетью (обратите внимание, как вы работаете, к примеру, с вашим любимым поисковиком), станет машинам не нужен.

На самом деле всё не так страшно и фантастично как кажется. Для того чтобы выделить имя или e-mail адрес средствами HTML были созданы тэги <NAME> и <ADDRESS>, но на их стандартизацию ушли годы, а многие из таких тэгов, содержащих «информацию об информации» так и не стали стандартом. Решение множества таких вот маленьких проблем и есть Semantic Web.

На данный момент вряд ли кто-то из людей, понимающих принципы работы семантической сети, сомневается в преимуществах этого решения или в его перспективах, но, к сожалению, сейчас не существует завершённой технологии разработки сервисов в рамках концепции семантической паутины. Перед разработчиками встаёт вопрос, изобретать что-то самостоятельно, или дождаться зрелости концепции и появления завершённых инструментов, а перед педагогами, обучать этому или нет, но чаще – как именно обучать.

Вопросы, мучающие разработчиков, в данной статье затронуты не будут, а вот попытка дать ответ на вопросы педагогов приведена ниже.

Семантическая паутина – технология молодая, не имеющая достаточного количества примеров успешной реализации, но существующие примеры заслуживают внимания. Чего стоит один только проект Goggle Apps [<https://www.google.com/a/>] или проект Wiki [<http://wiki.org/>]. Технология снабжения всех данных метаданными, упрощающими машинный анализ, безусловно, актуальна и востребована. Сейчас семантические сервисы предлагаются в основном крупными компаниями, более того, структура сервиса – интеллектуальная собственность этой компании и делиться ей никто не будет, но это только сейчас. В будущем, когда технология шагнёт в массы, рынку понадобятся специалисты, не только способные пользоваться семантическими сервисами, но и умеющие их создавать. На мысль о том, что технология шагнёт в массы, подталкивает и тот успех у пользователей, который сопровождает нынешние inside разработки крупных участников рынка, такие как сервисы Wiki или Google, и доходы этих крупных игроков.

Рынок поставщиков программного обеспечения откликается на перспективные потребности так же быстро, как и рынок поставщиков сервисов. Все без исключения игроки на рынке специализированного программного обеспечения предоставили инструменты разработки на языках, рекомендованных консорциумом W3 для реализации семантических сервисов. Это и технология Silverlight с собственной средой разработки от Microsoft, и технология AJAX, используемая и поддерживаемая сообществом, и многие другие технологии и среды разработки от компаний уровня Adobe и IBM. То есть, семантическая паутина на столько интересна потенциальным пользователям, что количество разработчиков, готовых к реализации сервисов так велико, что крупные поставщики программных продуктов готовы предоставить нужные инструменты для разработки.

Из всего вышесказанного можно сделать вывод, что изучать семантические технологии и обучать им надо. То есть, ответ на вопрос, надо ли изучать и обучать, утвердительный!

Литература

1. McDermott D. Critique of Pure Reason / Drew McDermott // Computational Intelligence. - 1987. - Vol. 3, Issue 1. - P. 151-160.
2. Gardner M. 1996. The universe in a handkerchief: Lewis Carroll's mathematical recreations, games, puzzles, and word plays / Martin Gardner. New York : Copernicus, 1996. - 158 p.
3. Weiser M. Some Computer Science Issues in Ubiquitous Computing / Mark Weiser // CACM. - 1993. - Vol. 36, Issue 7. - P. 74-83.