Альтернатива – Микроформаты для HTML

Web data that can be processed by machine.

Информация, которая получена из сети и которую можно машинно обработать. Впервые термит “Семантическая паутина” произнес Тимоти Джон Бернерс-Ли создатель World Wide Web и Глава коммитета всемирной паутины. Обычно страницы в интернете сделаны так, чтобы их было легко читать людям а не машинам. В сементической же сети информация может быль легко семантически (с сохранением смысла) интерпретирована машинами. The semantic web is a vision of information that can be readily interpreted by machines, so machines can perform more of the tedious work involved in finding, combining, and acting upon information on the web.

Взаимодействие различных онтологий Semantic Web tools allow them to link their schemes and translate their terms, gradually expanding the number of people and communities whose Web software can understand one another automatically.

Рассказать о существующем поиске по ключевым словам и о его недостатках.

Perhaps the most visible examples, though limited in scope, are the tagging systems that have flourished on the Web. These systems include del.icio.us, Digg and the DOI system used by publishers, as well as the sets of custom tags available on social sites such as MySpace and Flickr. In these schemes, people select common terms to describe information they find or post on certain Web sites. Those efforts, in turn, enable Web programs and browsers to find and crudely understand the tagged information—such as finding all Flickr photographs of sunrises and sunsets taken along the coast of the Pacific Ocean. Yet the tags within one system do not work on the other, even when the same term, such as “expensive,” is used. As a result, these systems cannot scale up to analyze all the information on the Web.

3 Практическая часть

Основная идея

Эта часть является описанием основных идей, на которых строится система ПИМ. Ключевая мысль заключается в том, чтобы перенести существующие технологии СВ на персональный компьютер. Ресурсы на компьютеры получают UNI адреса а их метадата конвертируется в RDF. Существующие десктопные приложения включены в эту структуру, а не заменены. Далее в деталях рассмотрено значене UNI в контексте его использования для идентифицирования идей и концепций.

**Semantic Desktop**

СВ позволяет улучшить WWW путем включения дополнительных метаданных. Этот же подход может быть использован и на персональных компьютерах. Подобно Вебу сервер приложений публикует доступную информацию. Локальный Semantic Web server может быть использован для того, чтобы получить доступ к информации расположенной на компьютере. Файлы и другие ресурсы могут быть использованы как ресурсы веб. Они идентифицируются с помощью UNI и позволяют создавать связи между различными документами. Существующая информация представлена в виде RDF документов и может быть использована любым другим приложением. Таким образом проявляется одно из основных свойств семантической паутины – различные программы одинаково обрабатывают данные с семантическими метаданными. Единый стандарт представления информации позволяет упростить взаимодействие между приложениями. В данном случает под приложениями понимаются десктопные программы, однако подобная структура может быть масштабирована до корпоративного уровня или даже глобального.

При использовании технологий семантической паутины на персональном компьютере в системной архитектуре создается семантический уровень. С этого момента ресурсы доступ к ресурсам может быть предоставлен посредством Semantic Web Server параллельно доступа к ним через уже установленные программы. Ресурсы идентифицируются независимо от формата своих файлов. На основе этого семантического уровня появляется возможность построить систему классификации, охватывающую все существующие ресурсы на компьютере. Такие концепции как «личное» или «работа» будучи один раз представлены в системе могут быть использованы всеми приложениями. Все ресурсы могут быть идентифицированы через UNI и просмотрены в браузеро-подобных программах. Поисковик, вроде Google, может быть создан поверх системы для того чтобы создать полнотекстовый индекс всех существующих текстовых данных. В отличие однонаправленных гипертекстовых ссылок в вебе связь между локальными ресурсами является двусторонней.

Отношение между двумя ресурсами может быть установлено в любой точке системы и быть доступно из любых других мест.

Как будет показано в примере привнесение технологий СВ в архитектуру персонального компьютера может быть с пользой использовано различными способами. При получения доступа к ресурсам через Semantic Web server открывает новые возможности по использованию существующих данных.

**Принципы**

Один компьютер – один пользователь. Представленная система семантической паутины спроектирована с учетом того, что ею будут пользоваться один пользователь, чьи ресурсы входят в эту семантическую сеть. Другие существующие PIM тоже поддерживают только одного пользователя, но не все. В данном случае ограничение на одного пользователя упрощают систему, так как отпадает необходимость поддерживать нескольких пользователей. Программа работает на одном персональном компьютере. Это не распределенная многопользовательская система и поэтому в ней отсутствуют механизмы пользовательской аутентификации и авторизации. В системе нет встроенной возможности обмена данными, если необходимо поделиться информацией с другими пользователями, она должна быть экспортирована и импортирована вручную. Вся RDF информация хранится на одном компьютере. Таким образом, весь RDF граф принадлежат одному пользователю. Так как каждый пользователь – уникальная личность и может быть идентифицирован, программа SemanticDesktop идентифицируется с помощью *hostname* (имя хоста). Так как имя хоста (доменное имя) глобально уникально, для каждого ресурса на компьютере возможно задать уникальный идентификатор. Если у компьютера нет уникального, доменного имени можно использовать другой способ. Этот способ заключается в том, чтобы использовать альтернативный идентификатор, который является доменным именем, принадлежащим владельцу компьютера плюс какое-нибудь имя для компьютера. Например, у моего компьютера нет уникального доменного имени, но я владею доменом

**http://www.arthur.lazy-magister.org**  
В этом случае уникальное имя для моего компьютера будет

**http://www.arthur.lazy-magister.org/learnmachine**

Следует отметить, что, хотя идентификатор и похож на URL, он всего лишь используется как уникальное имя и не указывает на какой-либо ресурс в сети.

Проблемы

На текущий момент не существует единой онтологии и списка ресурсов и свойств этих ресурсов. Dublin core и FOAF не обеспечивают необходимой универсальности. В будущем ожидается существование множества онтологий, которые в дальшейшем должны будут объединиться. Консорциум W3C рекомендует пользоваться существующими онтологиями а не изобретать свои.