

Ю. В. Малинина

**ЭЛЕКТРОННАЯ СРЕДА КОЛЛЕКТИВНОГО НАКОПЛЕНИЯ
И КАТАЛОГИЗАЦИИ ИНФОРМАЦИИ
ПО ПРЕОБРАЗОВАНИЯМ ПРОГРАММ***

ВВЕДЕНИЕ

Передача знаний происходит во время общения между людьми, направленного на получение необходимых знаний для решения задач или принятия решений. Коммуникации могут быть персональными и групповыми, непосредственными или заочными. Персональные коммуникации (общение) используются в повседневной жизни, например, тогда, когда человек обращается за советом или консультацией к коллеге. Для обеспечения эффективности групповых коммуникаций применяются различные мероприятия, организующие процесс общения в группах. К таким мероприятиям можно отнести привычные совещания, семинары, конференции, съезды и т.д. Заочные коммуникации обычно происходят, например, посредством бумажных или электронных документов и писем.

Между учеными до сих пор нет единого мнения о том, что такое теория коммуникации и что должно быть предметом ее исследования. Профессор Роберт Т. Крейг [2] считает, что существование разных трактовок понятия «коммуникация» связано с тем, что исследованиями в этой области занимаются специалисты разных профессий, каждый из которых рассматривает коммуникации применительно к своему роду деятельности и зачастую не интересуется работами коллег из других сфер. Однако уже не вызывает сомнения то, что Интернет и сетевые технологии могут существенно влиять на процесс коммуникаций. Особое внимание привлекают различные способы асинхронного взаимодействия, которые снимают трудности, связанные с организацией групповых и заочных коммуникаций. Если раньше они были ограничены текстовой формой электронных писем и бюллетеней, то теперь эта форма взаимодействия стала более гибкой.

Если обратиться к творческому наследию Герберта Маршалла Маклюэна, то в своих монографиях он утверждал, что западная цивилизация дос-

* Работа выполнена при финансовой поддержке Научной программы «Университеты России» (грант № УР.04.01.027) и Министерства образования РФ (грант № Е02-1.0-42).

тигла «водораздела» в XX веке, имеющего не меньшее значение, чем эпоха Ренессанса. Качественные сдвиги в истории человечества, согласно Маклюэну [3], связаны с появлением новых технических средств коммуникации. Отсюда и следует его знаменитое высказывание о том, что "средство и есть содержание" ("The Medium is the Message" — сообщением, передаваемым средством общения, является само это средство). Он говорит, что попытки выполнять сегодняшние обязанности с помощью вчерашних орудий — вчерашних концепций, где информация скудна, но упорядочена и построена по отдельным классифицированным планам, темам и графикам, уже не отвечают современным требованиям. "Технологии — это не просто изобретения, которые используют люди, это средства, с помощью которых люди создаются заново... Будущее — адрес нашего проживания".

Согласно Нельсону [4], одному из пионеров гипертекстов, реальная структура документов и источников игнорируется существующими сегодня форматами, имитирующими бумажные документы. Пересечение и параллелизм документов, ключевыми примерами которых могут служить аннотации и взаимные связи, различные схемы цитирования и доступность оригинала цитируемого текста, практически не используются в повседневной работе.

Сегодня можно заметить, что компьютерные науки ставят в центр своих изысканий два традиционных конструкта — имитацию иерархии и имитацию бумаги. Существует популярный миф, что «структура» всегда означает иерархию. Существует также популярная концепция, что электронный документ должен имитировать документ бумажный.

Зачастую информационная структура не может быть верно представлена иерархией. Такие взаимосвязи, как параллелизм, перекрестные связи, взаимное проникновение и взаимное присутствие (присутствие одного элемента в нескольких местах), теряются при построении иерархии, когда делается выбор в пользу только одного типа отношений. Например, в представлении файловых систем вы можете выбрать только один тип иерархии и расположить объекты по шкале времени или важности в ущерб всем остальным типам отношений. Эти отношения становятся основными и определяют все последующие соглашения — псевдонимы, сокращения и базы данных.

Как следствие, суперсовременные системы работы с документами не соответствуют тому гигантскому потенциалу, которым обладает компьютер, не используют и малой толики из возможностей, которые открываются в случае, если документ представлен в форме, читаемой машиной. Вот и вы-

ходит, что, в конечном счете, компьютер на столе служит не более чем удобной пишущей машинкой и компактным, безопасным архивом.

Однако если прислушаться к Нельсону и покинуть привычное для нас поле традиционных структур, перед нами открываются многочисленные альтернативные возможности. Рассмотрим некоторые из них.

НОВЫЕ ВОЗМОЖНОСТИ

Тот факт, что научная деятельность и творчество уже не возможны без существенных изменений в способах коммуникации и способах размышления, отмечался уже в середине прошлого столетия в работах основателей кибернетики. Еще Ваннавер Буш описывал новый способ мышления при помощи машинных средств обработки и передачи информации [1]. Валентин Турчин рассматривает творчество как метасистемный переход и также полагает, что использование машин позволит перевести мыслительную деятельность на более высокий уровень [5]. Новые цифровые технологии могут освободить наш мозг от механических аспектов мышления для решения творческих задач.

WikiWiki — коллективный гипертекст

Не секрет, что процесс мышления обладает определенной спецификой: идеи развиваются на разных уровнях, в разных направлениях одновременно и взаимосвязано. Записывать такую систему идей и потом читать ее достаточно трудно. Традиционные плоские текстовые файлы связывают нас необходимостью писать и читать параграфы текста, главным образом, в линейной форме. В традиционной книжной форме представления текста присутствуют элементы нелинейности: комментарии в скобках, комментарии в сносках, ссылки на предыдущие или последующие разделы, ссылки на литературу, фрагменты текста, выделенные другим шрифтом, которые позволяют автору неявно сказать: «Это не входит в основной текст, но имеет к нему отношение; если Вам интересно, можете посмотреть». Кроме того, существует много стилистических оборотов, позволяющих представить нелинейный по существу ход рассуждений в последовательной линейной форме. Однако этих форм нелинейности не всегда бывает достаточно, не всегда они удобны при просмотре текста, особенно на компьютере.

Известно, что существующие тексты и без компьютерных сетей образуют единое пространство [6]. Но при этом система адресации имеет достаточно запутанный характер, поскольку хранится частично в книжной, а час-

тично — в человеческой памяти. В действительности библиотечные и информационные науки в основном и состоят из исследования возможностей ссылок. Любой, кто проводил исследования, понимает, что значительную часть работы составляет получение нужной литературы, просмотр ссылок на литературу, просмотр терминов в справочниках или глоссариях, проверка таблиц или рисунков, анализ оглавлений или перечислений, составление собственных пометок. Даже при простом чтении постоянно встречаются ссылки на другие разделы или главы, подстрочные комментарии, литературу, примеры, таблицы, рисунки, приложения. Часто некоторые главы пропускаются читателем, которого не интересуют технические подробности. Однако перед традиционной литературой стоят определенные проблемы.

- Большинство ссылок не помогают вернуться назад: читателю бывает очень сложно найти те книги или статьи, которые ссылаются на ту статью, которую он читает. И электронные варианты публикаций не являются исключением, а с завидным упорством повторяют этот подход.
- Читатель должен размещать свои пометки либо на полях книги или статьи, либо помещать их на отдельном листе. Подобный подход обычно применяется и при работе с файлами.
- Наконец, следование по ссылкам в бумажных документах (или их электронных аналогах) требует постоянной затраты сил, внимания, дает задержки, даже если читатель работает в хорошо оборудованной библиотеке.

Обратим свое внимание на ставший уже привычным при нынешнем развитии Интернета – гипертекст и попробуем посмотреть на него с другой стороны.

Сама концепция гипертекста достаточно проста. Однако несмотря на свою простоту, эта идея дала толчок многим интересным идеям и разработкам.

Фактически, гипертекст — это множество документов, связанных между собой ссылками. На первый взгляд, основная суть гипертекста состоит в том, чтобы предоставлять быстрый и удобный интерфейс к текстовым фрагментам. Но такое описание ничего не говорит об его реальной роли в развитии коммуникаций. Однако, если рассмотреть гипертекст как специальную, поддерживаемую компьютером среду для мышления и общения, можно сделать ряд интересных наблюдений.

Наиболее прогрессивная и радикальная модель коллективного гипертекста — WikiWiki. Тексты WikiWiki связаны между собой при помощи самой простой системы адресации. Само слово — название ресурса — и является его адресом в базе данных. Для того чтобы сделать ссылку на ресурс, достаточно просто упомянуть это слово в тексте.

Гипертекст (WikiWiki) позволяет автору делать подобные ссылки, а читателю самому выбирать, по каким связям он будет идти при просмотре данного текста. С этой точки зрения, WikiWiki упрощает работу мыслителя и автора: не требуется постоянно решать — вводить ли данную мысль в основной поток изложения или нет. В этом смысле «ссылочность» является основой мощности гипертекстов: именно ссылки, поддерживаемые компьютером, расширяют текст за пределы одномерного потока.

В то же время, если гипертекст используется как инструмент для мышления, написания, проектирования, может возникнуть естественное соответствие между объектами окружающего мира и узлами в базе данных гипертекста. Пользуясь достоинствами такого объектно-ориентированного аспекта, при помощи гипертекста можно построить гибкую сеть, которая будет моделировать некоторую проблему (или ее решение). В таких приложениях связи могут быть даже менее важными, чем узлы: связи формируют смысловые группы, которые собирают узлы вместе, но основной акцент делается именно на узлы.

WikiWiki позволяет легко использовать в повседневной деятельности все преимущества, предоставляемые гипертекстами.

- *Простота следования по ссылкам:* поддержка компьютером при следовании по ссылкам позволяет одинаково легко следовать как вперед по ссылкам, так и возвращаться назад.
- *Простота создания новых ссылок:* пользователь может развивать свою сеть или просто комментировать чей-то документ.
- *Структуризация информации:* к неструктурированной информации можно применять как иерархический способ организации, так и не-иерархический; более того, можно на одном и том же материале организовать несколько разных иерархий.
- *Глобальный взгляд:* специальные системы просмотра могут обеспечить глобальный взгляд на документ как на сеть узлов, что существенно для очень больших или сложных документов.

- *Текстовые узлы* могут быть собраны вместе различными способами, давая возможность одному и тому же документу выполнять различные функции.
- *Модульность информации*: так как на один и тот же текстовый сегмент можно ссылаться из нескольких мест, мысли могут быть выражены с меньшими перекрытиями и дублированием.
- *Связность информации*: ссылки становятся неотъемлемой частью текста, и если какая-то часть текста переносится в другое место, даже в другой документ, информационные ссылки продолжают давать прямой доступ к данному фрагменту текста.
- *Взаимодействие с использованием документа*: то, что пользователь сам выбирает путь, по которому он просматривает гипертекст, делает пользователя активным участником процесса соотнесения гипертекстового документа к задаче, стоящей перед пользователем.
- *Совместная работа*: возможность совместной работы нескольких авторов над одним документом, что дает принципиально новые возможности.

Конечно, при увеличении объема и сложности информации, представляемой в виде гипертекста, зачастую читателю все труднее становится понимать, в каком месте сети он сейчас находится и как ему перейти в другое нужное место гипертекста. Эта проблема возникает из-за того, что в гипертексте, в отличие от линейного текста, у пользователя гораздо больше возможных путей просмотра, больше степеней свободы.

Существуют два основных технических решения, помогающих решать эту проблему: подсистема графического просмотра гипертекстовой сети и механизм поиска по запросу. Подсистема графического просмотра сети — это подсистема, позволяющая в явном виде представить структуру гипертекста в виде сетевого графа и показать положение пользователя в этой сети. Естественно, такие подсистемы требуют очень высокой разрешающей способности компьютерных дисплеев. Кроме того, для того чтобы разместить узлы и связи в двух- или трехмерном пространстве, придать им полезные визуальные особенности (цвет, форму, размер, фактуру), придерживаться полезных соглашений (не размещать два узла слишком близко друг от друга и т.п.), разработчики таких подсистем должны решить очень сложную программистскую задачу. Но даже и такое сложное программирование все равно оказывается стоящим перед тем фактом, что для сложного ин-

формационного пространства просто не существует естественной топологии.

Другое решение этой проблемы — применить типовую технику поиска по запросу из баз данных. Запрос на поиск обычно организуется в виде некоторой логической комбинации поиска по ключевому слову, по подстроке, по свойствам узлов или связей.

Создание документа в WikiWiki

Работа авторов документов обычно рассматривается как работа на уровне слов и предложений. Ясно, что текстовый редактор — это хороший инструмент для авторов. Если же рассматривать работу по написанию текстов как работу по структуризации идей, упорядочиванию представлений, использованию понятий, текстового редактора становится недостаточно. Очень немногие авторы просто садятся и пишут законченный текст. В каком-то смысле авторы — это проектировщики текстовых документов. И единицей такого уровня авторской работы является мысль или понятие, и работа такого уровня может быть легко поддержана при помощи гипертекстов. По мере того как автор приходит к новым мыслям, он может развивать их в виде узлов и затем привязывать эти мысли к единому целому или делать их изолированными. Таким образом, гипертексты помогают автору двигаться от слабо структурированной сети мыслей к законченному документу, годному для печати.

WikiWiki вводит иной подход создания Интернет страниц. Обычно требуется, чтобы сначала была создана страница, а уже затем на эту страницу была бы сделана ссылка. В WikiWiki в таком предварительном создании нет необходимости. Всякое новое определение сначала вводится в тексте, а потом уже может быть разъяснено. Мы можем ввести в тексте новые термины, например, так называемое WikiWord-слово. После сохранения текста мы получаем страничку, на которой после слова-определения WikiWord появилась ссылка со знаком вопроса, сигнализирующая о том, что определение этого термина еще не задано. Нажав на эту ссылку, мы создадим новую страницу, которая будет содержать определение понятия WikiWord.

Преимущество WikiWiki технологии заключается в том, что она позволяет реализовать коллективную работу с гипертекстовыми документами с использованием простого языка разметки и обычного браузера. Например, должно быть выработано общее определение и представление о понятии «преобразование программы». В WikiWiki нельзя завести несколько объектов ПреобразованиеПрограммы. С другой стороны, к правке понятия Пре-

образовании. Программы имеют доступ все участники, и таким образом реализуется возможность хранить несколько точек зрения на некоторое понятие в одном ресурсе, которая особенно важна в случаях, когда в рамках совместной работы формируются и обсуждаются термины и определения.

На сегодня существует много различных реализаций WikiWiki; их объединяет то, что WikiWiki-страницы имеют два представления. Одно — это вид, который требуется для вывода (обычно это HTML), а второе, доступное для пользователей, пишется на упрощенном языке разметки, незначительно отличающемся от естественного языка, синтаксис и стиль которого могут варьироваться в зависимости от конкретной реализации. Система WikiWiki дает удобные возможности при создании сложных источников информации. Линейный (не гипертекстовый) документ нормально может читаться только единственным способом — каким он организован. У нелинейного же текста есть существенное достоинство: такой текст можно организовать различными способами, в зависимости от точек зрения.

Идея WikiWiki хорошо воспринята целым рядом технологических сообществ, но подобные среды оставались атрибутами этих сообществ, не выходя вонне, размеры баз данных в них не превышали нескольких сотен страниц. Настоящий прорыв произошел в 2001 г., когда Лари Сандр дал старт проекту WikiWikipedia. На октябрь 2003 г. число статей в WikiWikipedia превышает 150 тыс.

Ссылки по теме:

- <http://c2.com/cgi/wiki/wiki> — самый первый WikiWiki. — WikiWiki история развивалась как техническая поддержка хранилища образов и моделей (Patterns).
- <http://www.WikiWikipedia.org/> — сайт-энциклопедия в стиле WikiWiki. (<http://ru.WikiWikipedia.org/>)

К несомненным достоинствам WikiWiki относится то, что этот продукт можно использовать и в индивидуальной, и в коллективной практике. Для индивидуального использования достаточно на личном компьютере установить Интернет-сервер и какой-нибудь из клонов WikiWiki (возможно понадобится установка Perl, php, java — в зависимости от выбранного клона), и все это будет быстро и хорошо работать. WikiWiki является хорошим инструментом для хранения записок по разным темам.

Обычно WikiWiki поддерживает следующие возможности редактирования.

- Выделение текста.
- Создание списочных структур.
- Простое форматирование текста.
- Автоматическое задание ссылки для слов определенного формата.
- Явное указание ссылок или их отмена.
- Использование некоторых HTML тэгов.

Интересного эффекта можно добиться интеграцией WikiWiki с коллекцией. В результате этого симбиоза к привычной для всякой WikiWiki возможности создания и редактирования гипертекста добавилась возможность напрямую ссылаться на не редактируемые тексты, собранные в коллекции. Ссылки на авторов коллекции ведут к их текстам. К авторским текстам добавилась возможность «кликнуть» имя автора текста и получить страницы WikiWiki, на которых происходит обсуждение и использование этого текста (своего рода заметки на полях электронных текстов в стиле WikiWiki).

Основные возможности, предоставляемые WikiWiki.

- Относительная простота «входа» в систему.
- Возможность редактирования любой страницы системы любым пользователем.
- Возможность многопользовательской работы над документами.
- При многопользовательской работе не надо строго формализовать работу группы.
- Контроль версий документов.
- Возможность проставления ссылок на несуществующие страницы, а также возможность проставления ссылок по названию страниц, а не по адресу.
- Упрощенная разметка текста: это не WYSIWYG, но все же и не HTML. Впрочем, это можно отнести и к минусам, зависит от того, как вы привыкли редактировать.

WikiWiki как хранилище групповой памяти

Интерес к WikiWiki связан с поиском средств, которые бы поддерживали коллективную научную деятельность. Идея построения направленного графа из неформальных текстовых элементов близка идее семантической сети из области искусственного интеллекта. Семантическая сеть — это схема представления знаний, состоящая из направленного графа, в котором понятия представляются как узлы, а отношения между этими понятиями — как связи между узлами.

Семантическую сеть отличает то, что понятия индексируются посредством их семантического содержания, а не каким-либо внешним (возможно, алфавитным) порядком. Одно из преимуществ семантических сетей перед другими схемами представления знаний заключается в том, что использование их естественно, т.к. близкие понятия стремятся связаться в единый пучок сети. Соответственно, легко распознать неполные или неадекватные понятия, так как смысловой контекст определяется именно связями с соседними понятиями.

Здесь можно увидеть прямую аналогию с гипертекстами, и в частности с WikiWiki технологией. Отличие состоит в том, что инженер знаний стремится построить такие представления, которые могли бы автоматически интерпретироваться, в то время как цель автора гипертекста — собрать взаимосвязанную систему мыслей, не претендуя на машинную интерпретацию.

ЗАКЛЮЧЕНИЕ

Итак, WikiWiki — не случайное событие, а результат закономерного эволюционного процесса развития гипертекстов. Использование WikiWiki в качестве инструмента для организации среды построения каталога по преобразованиям программ, на наш взгляд, позволит найти дополнительные взаимосвязи в области преобразований программ и облегчить повседневную работу специалистов в этой области.

WikiWiki позволяет организовать в сети для всех участников коллективной деятельности общее равнодоступное пространство, где они могут размещать свои мысли в форме письменных высказываний, связывать эти высказывания между собой, редактировать и корректировать свои и чужие высказывания. Такую электронную среду можно рассматривать как аналог доски, на которой может писать и рисовать каждый. Электронная среда имеет очевидное преимущество, поскольку неограниченные размеры позволяют хранить всю историю взаимодействия и все разнообразие связей. Публикуемые в сети документы наряду с текстом могут содержать ссылки на графику, диаграммы, звуковые файлы и другие связанные ресурсы. Общая деятельность может быть представлена в форме, облегчающей индивидуальное восприятие.

СПИСОК ЛИТЕРАТУРЫ

1. Bush V. As we may think // The Atlantic Monthly. —1945. — Vol. 176(1). — P.101–108 — <http://www.csi.uottawa.ca/~dduchier/misc/vbush/awmt.html>
2. Craig R.T. Communication Theory as a Field // Communication Theory. —1999. — N 9 — P. 217–242 — <http://www.comm.cornell.edu/comm680/craig.pdf>.
3. McLuhan M. The Medium is the Message. — NY, 1967. — <http://uic.nnov.ru/pustyn/lib/macLU.ru.html>
4. Nelson T. Structure, traditions and possibility // Proc. of the 14th ACM Conf. on Hypertext and hypermedia, Nottingham, UK, August 26–30, 2003. — ACM Press, 2003. — P. 1–1
5. Турчин В. Феномен науки. Кибернетический подход к эволюции — М: ЭТС, 2000. — 368 с. — <http://www.refal.ru/turchin/phenomenon/index.htm>
6. Фуко М. Археология знания. — Киев: Ника-центр, 1996. — 208с. — http://yanko.lib.ru/books/cultur/foucalt_larcheologie_du_savoir_ru.htm