

УДК 004.72:004.414.23

А.М. Кашевник (4 курс, каф. РВКС), А.В. Смирнов, д.т.н., проф. СПИИРАН

ФОРМИРОВАНИЕ ПРОФИЛЯ ПОЛЬЗОВАТЕЛЯ В РАСПРЕДЕЛЁННОЙ СИСТЕМЕ ИНТЕГРАЦИИ ЗНАНИЙ

Быстрое развитие Интернет позволило огромному объёму информации о различных проблемных областях стать доступным для пользователей. Поскольку эта информация представлена в разных форматах и различными средствами, проблемы совместимости форматов, реализации таких средств как распознавание и интеграция знаний из распределённых источников является актуальной. Возникла необходимость разработки подхода для быстрого слияния знаний из распределённых источников, что необходимо для глобального понимания происходящих процессов и явлений, динамического планирования и глобального обмена знаниями.

Процесс извлечения, интеграции и доставки нужного знания из подходящего источника в правильном контексте заданному пользователю в нужное время для определённых целей называется “Логистикой Знаний” (Смирнов А., Пашкин М., Шилов Н., Левашова Т. Agent-Based Knowledge Fusion in Scalable Information Environment: Major Principles and System Framework).

Решать проблему логистики знаний можно путём интегрирования знаний из различных источников (интеграции знаний).

Для системы интеграции знаний важно знать информацию о пользователях, с которыми она работает (профиль пользователя). Наблюдая за пользователями, изучая их интересы и объединяя их в группы по можно подготовить шаблоны, по которым пользователь может быстро наглядно и чётко получить ответ на интересующую его тему.

Пользовательский профиль можно разделить на три типа данных: персональные данные, системные данные, данные, полученные в процессе работы. Персональные данные содержат личные данные о пользователе. Если два человека живут в одном месте, занимаются схожими вещами, то с высокой долей вероятности можно сказать, что у них будут схожие интересы, и их можно объединить в группу пользователей по интересам. Системные данные определяют пользователя как потребитель знаний, эксперт и соответственно предпочтения пользователя. Если системе не удастся найти знание по запросу пользователя, то она обращается за помощью к эксперту. Данные, полученные в процессе работы. Система «следит» за пользователем: в базе данных хранится история всех запросов пользователя, посещений пользователем системы.

Планируется внедрение технологии Machine Learning. То есть надо активно использовать данные, полученные в результате работы системы.

Достоинства системы:

- Работа с базами данных осуществляется через ODBC: может быть подключена любая БД имеющая ODBC драйвера для типов СУБД.
- Отсутствие специального ПО у пользователя: PHP-скрипты выполняются на сервере, на компьютере пользователя – HTML код.