

Сервисы Explain платформы ZooSPACE и адаптивные пользовательские интерфейсы

© О. Л. Жижимов

Институт вычислительных технологий СО РАН,
Новосибирск
zhizhim@mail.ru

Аннотация

Рассмотрена технология реализации сервисов Explain на платформе ZooSPACE. Эта платформа предназначена для обеспечения доступа к распределенным разнородным информационным ресурсам по протоколам SRW/SRU, Z39.50 в том числе посредством графических пользовательских WEB-интерфейсов. Обсуждаются детали реализации пользовательских интерфейсов для поиска информации в разнородной информационной среде, адаптируемых к структуре и функциональности информационных источников. Приводятся примеры реализации для распределенной информационной системы ИВТ СО РАН.

Работа выполнена при финансовой поддержке Министерства образования и науки Российской Федерации (ГК № 07.514.11.4130, ГК № 14.521.11.0004), при частичной поддержке РФФИ (проект 13-07-00859), интеграционных проектов СО РАН, проектов ФНИ (IV.38.1.1) и президентской программы «Ведущие научные школы РФ» (грант НШ–5006.2014.9)

1 Платформа ZooSPACE

Платформа ZooSPACE была разработана в рамках выполнения Государственного контракта № 07.514.11.4130 от 6.06.2012 Министерства образования и науки РФ в рамках ФЦП «Исследования и разработки по приоритетным направлениям развития научно-технологического комплекса России на 2007-2013 годы» по теме «Разработка принципов и программных средств виртуальной интеграции распределённых источников данных на основе международных стандартов для создания масштабных информационных инфраструктур», одним из результатов выполнения которого явилось создание технологической платформы интеграции разнородных распределенных данных ZooSPACE

[1-3]. Платформа ZooSPACE ориентирована на создание распределенных информационных систем (РИС), интегрирующих разнородные информационные ресурсы, управляемые различными СУБД, на основе единых политик организации доступа к этим ресурсам. Программные компоненты ZooSPACE функционируют на различных программно-аппаратных платформах серверов, сосредоточенных в узлах распределенной системы. Взаимодействие узлов между собой осуществляется посредством сетевых протоколов прикладного уровня на основе транспортного протокола TCP/IP. Количество узлов в ZooSPACE не нормируется и может быть любым. Система ZooSPACE может состоять из одного единственного узла. Такой выбор инфраструктуры узлов позволяет обеспечить достаточно гибкую распределенную информационную систему и реализовать всю необходимую функциональность, которая обеспечивается подсистемами ZooSPACE.

В качестве подсистем ZooSPACE выступают следующие:

ZooSPACE-L обеспечивает функционирование справочной и административной подсистемы ZooSPACE. Она интегрирует совокупность LDAP серверов узлов, функционирующих в соответствии с единой для всех политикой и хранящих в виде единой иерархической базы данных всю конфигурационную и административную информацию ZooSPACE.

ZooSPACE-Z обеспечивает функционирование подсистемы доступа к базам данных системы ZooSPACE. Она интегрирует совокупность Z39.50 и SRW/SRU серверов узлов, функционирующих в соответствии с единой для всех политикой. В качестве серверов Z39.50 и SRW/SRU используется модифицированный сервер ZooPARK – ZooPARK-ZS.

ZooSPACE-M обеспечивает функционирование системы мониторинга всех компонент ZooSPACE.

Труды XVII Международной конференции DAMDID/RCDL'2015 «Аналитика и управление данными в областях с интенсивным использованием данных», Обнинск, 13-16 октября 2015

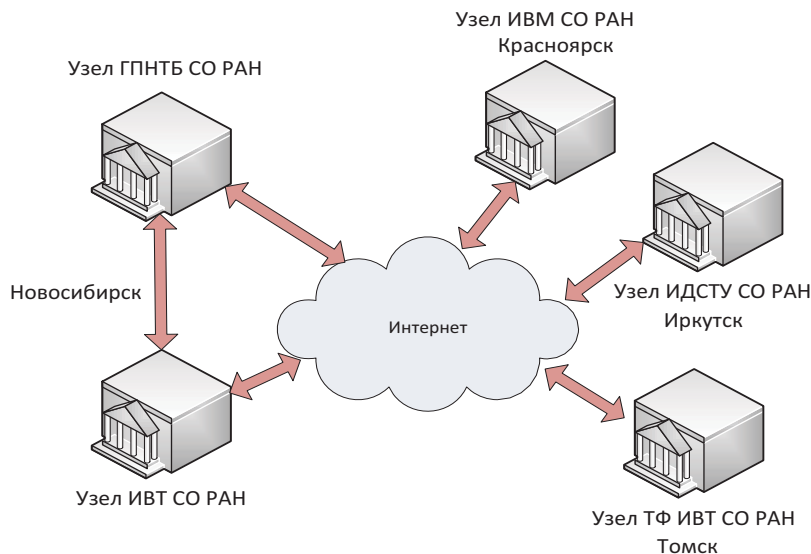


Рис. 1. Инфраструктура узлов РИС СО РАН на основе ZooSPACE

ZooSPACE-S обеспечивает функционирование подсистемы сбора статистики работы всех компонент ZooSPACE.

ZooSPACE-W предназначена для предоставления административных и пользовательских WEB-интерфейсов для доступа к ZooSPACE. Программное обеспечение сервера подсистемы ZooSPACE-W состоит из блоков:

- **Блок Z** реализует интерфейсы доступа к подсистеме ZooSPACE-Z. Этот блок обеспечивает поиск и представление данных из различных СУБД в соответствии с выбранным профилем.
- **Блок L** реализует интерфейсы доступа к подсистеме ZooSPACE-L.
- **Блок S** реализует интерфейсы доступа к подсистеме ZooSPACE-S.
- **Блок M** реализует интерфейсы доступа к подсистеме ZooSPACE-M.

В процессе выполнения различных проектов, в том числе и интеграционных проектов СО РАН, экспериментальная РИС на основе ZooSPACE, созданная как экспериментальный стенд для выполнения Государственного контракта № 07.514.11.4130, была существенно расширена как в части количества узлов, так и в части номенклатуры интегрируемых информационных ресурсов.

На сегодняшний день РИС ZooSPACE включает пять узлов (ИВТ СО РАН, ГПНТБ СО РАН, ТФ ИВТ СО РАН, ИВМ СО РАН, ИДСТУ СО РАН) в Новосибирске, Томске, Красноярске и Иркутске (см. рис. 1). Каждый узел содержит серверы LDAP, Z39.50/SRW/SRU (ZooPARK-ZS), WEB-сервер, что позволяет организовать доступ к различным СУБД и предоставить унифицированные пользовательские и

административные интерфейсы для доступа к информационным ресурсам. На текущий момент система обеспечивает доступ к более 80 базам данных различной тематической направленности [3]. Номенклатура доступных в ZooSPACE информационных ресурсов постоянно расширяется.

2 Сервисы Explain

Разнородность информационных ресурсов, интегрируемых платформой ZooSPACE, требует наличия специальных сервисов для аннотирования функциональности и структуры каждого информационного источника в соответствии с едиными для всей системы правилами (см., например, [4-5]). В качестве таких правил платформа ZooSPACE использует спецификации Explain для протоколов Z39.50 [6] и ZeeRex [7-8] для протоколов SRW/SRU [9]. Следует также отметить, что информационные ресурсы, интегрированные в ZooSPACE посредством серверов ZooPARK-ZS автоматически доступны как по протоколу Z39.50, так и по протоколам SRW и SRU.

При генерации информации Explain используется тот факт, что основная ее часть так или иначе уже содержится в подсистеме ZooSPACE-L и доступна по протоколу LDAP. Поэтому часть информации Explain извлекается из каталога ZooSPACE-L в момент старта сервера ZooPARK-ZS, помещается в хранилище и индексируется при помощи Zebra [10], входящей в состав одной из подсистем сервера ZooPARK-ZS (см. рис.2). Получаемая таким образом база данных с именем IR-Explain-1 полностью соответствует спецификациям Z39.50 и содержит информацию в категориях: CategoryList, TargetInfo, DataBaseInfo, AttributeDetails.

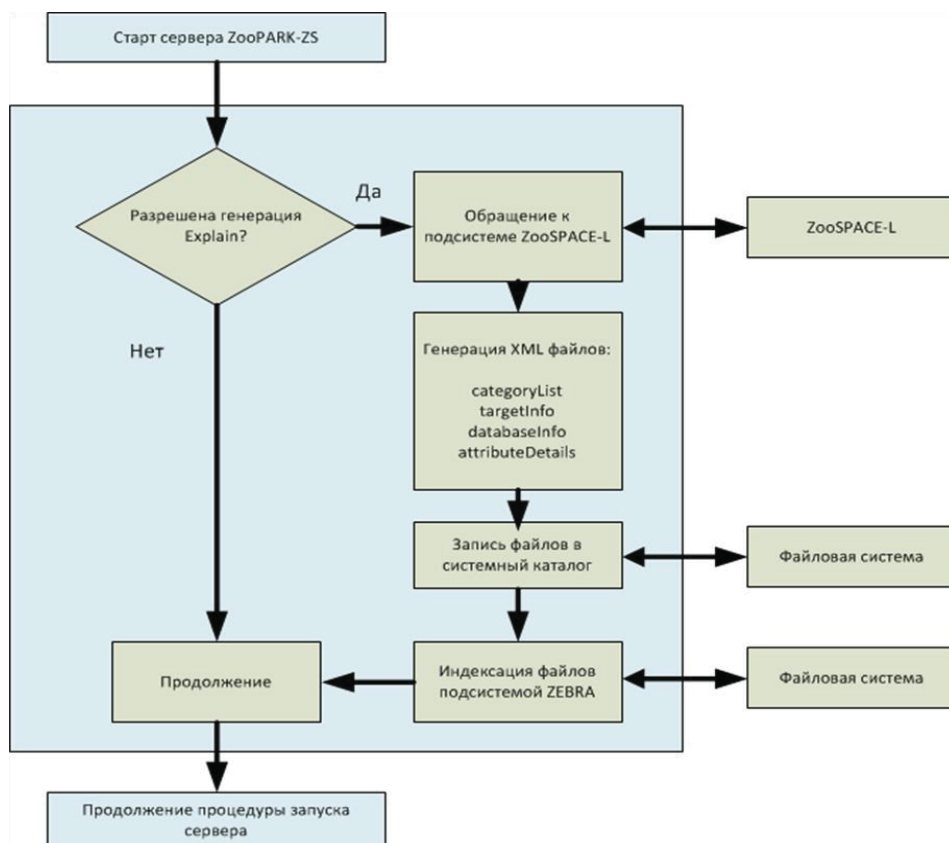


Рис. 2. Генерация статической информации Explain

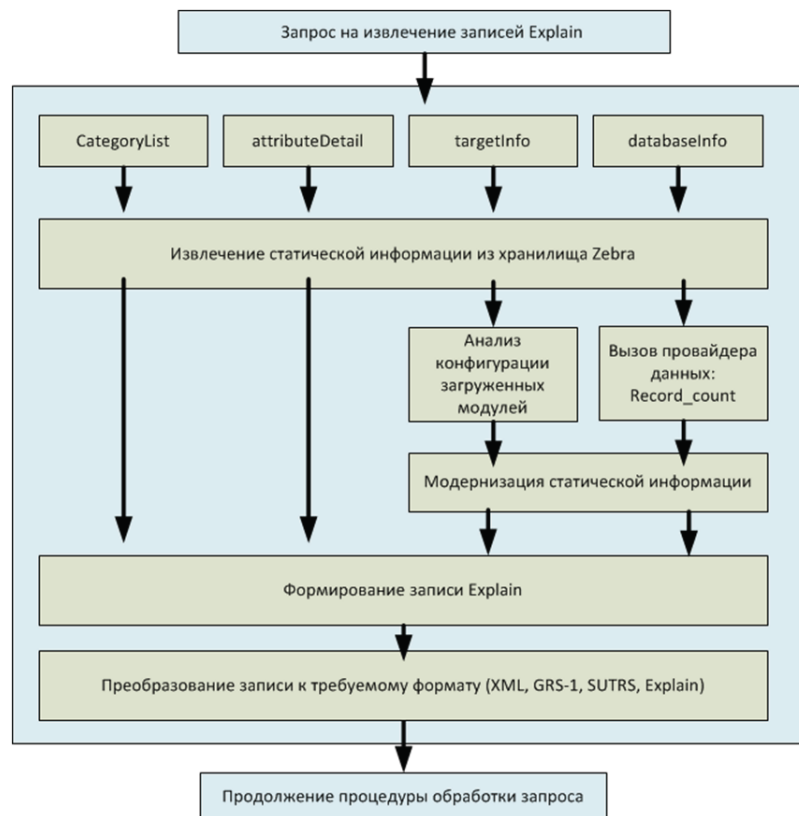


Рис. 3. Генерация динамической информации Explain

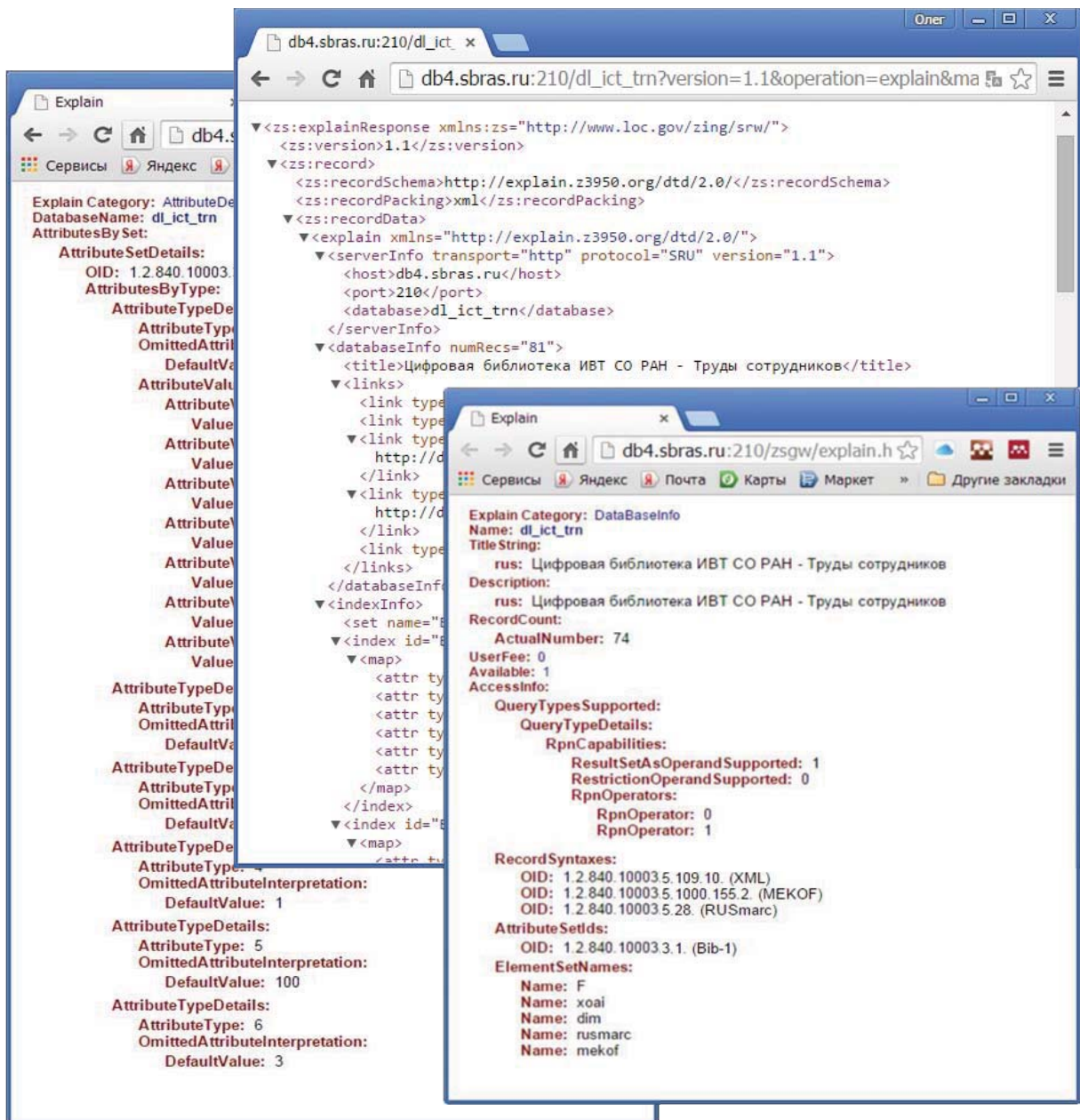


Рис. 4. Представление записи Explain в схеме ZeeRex, а также представление записей IR-Explain-1 в категориях DataBaseInfo, AttributeDetails (Z39.50)

Однако, при формировании записей Explain для пользователя статической информации, полученной из каталога ZooSPACE-L, явно недостаточно. Например, в ней заведомо отсутствует информация о текущем количестве записей в базах данных и о доступности баз данных. Для включения этой информации в результирующие записи Explain в момент их формирования для пользователя происходит вызов соответствующего провайдера данных для получения этой информации в соответствии с алгоритмом, представленном на рис. 3.

Информация Explain представляется в соответствии со спецификациями Z39.50 в форматах Explain, GRS-1, XML, SUTRS и в соответствии со спецификациями SRW/SRU в формате XML в схеме ZeeRex (рис.4).

При этом записи Explain в схеме ZeeRex формируются только динамически в момент соответствующего запроса, например

```
http://db4.sbras.ru:210/dl_ict_trn?
version=1.1&operation=explain
&maximumRecords=1&startRecord=1
&recordPacking=xml
```

3 Адаптивные пользовательские интерфейсы

Под адаптивными пользовательскими интерфейсами обычно понимается их способность подстраиваться под потребности пользователя.

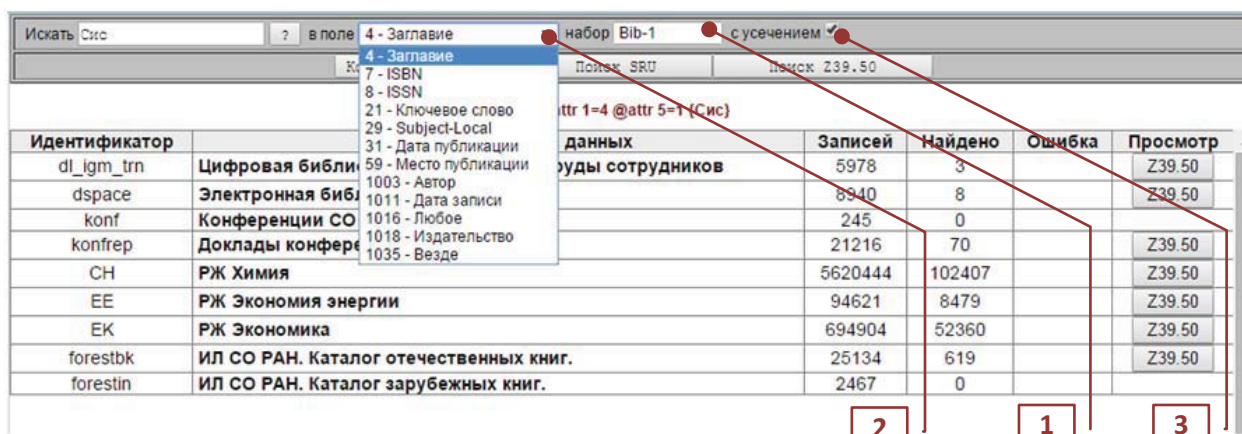


Рис. 5. Адаптивные элементы управления для поиска информации



Рис. 6. Адаптивные элементы управления для просмотра информации

Однако существует и другая сторона адаптивности - способность интерфейсов подстраиваться под структуру и функциональность информационных источников. Именно такая адаптивность будет рассматриваться ниже.

Пользовательские интерфейсы для работы с информационными ресурсами должны, как минимум, включать элементы управления поиском информации и способом отображения извлекаемой информации. Эта задача просто решается для однородных источников информации, однако, для неоднородных источников информации возникают проблемы. С одной стороны, элементы управления должны быть настроены на наиболее широкие диапазоны выбора параметров поиска и представления информации, но, с другой стороны, каждому конкретному информационному ресурсу как правило соответствует лишь достаточно узкий спектр параметров из потенциально возможных их значений. Для исключения перегруженности пользовательских интерфейсов возможностью применения недопустимых значений при выборе параметров элементов управления для различных информационных ресурсов требуется привлечение дополнительной информации о конкретном информационном ресурсе.

Пользовательские интерфейсы, подстраивающиеся под характеристики информационных ресурсов, ниже будут называться *адаптивными*.

Для реализации адаптивных в указанном смысле пользовательских интерфейсов необходима входная информация о функциональных свойствах каждого источника данных. Такую информацию при использовании Z39.50 и SRW/SRU предоставляют

сервисы Explain [6-8] в различных своих модификациях. Именно эти сервисы используются в ZooSPACE.

Адаптивность пользовательских интерфейсов в подсистеме ZooSPACE-W реализована для поиска и представления информации для протоколов Z39.50 и SRU. Адаптация происходит на стороне WEB-сервера после выбора списка информационных источников на основе информации Explain (см. выше).

На основе сформированного пользователем списка информационных источников WEB-сервер формирует для каждого источника Explain-карту, обрабатывая результаты соответствующих запросов к источнику в соответствии с действующими протокольными процедурами. В частности, для SRW/SRU/ZeeRex результатом explain запроса является xml структура в схеме ZeeRex (см. выше), которая легко обрабатывается при помощи, например, XSLT. Сформированные данные сохраняются в сессионных переменных на стороне сервера для дальнейшего использования при обработке AJAX запросов клиента.

Для клиента при формировании интерфейсов для поиска информации (см. рис. 5) учитывается для каждого конкретного ресурса (запросы RPN-1 или PQF):

- Список наборов поисковых атрибутов (метка 1 на рис.5)
- Список поддерживаемых поисковых атрибутов для каждого набора атрибутов (метка 2 на рис.6)
- Возможные комбинации поисковых атрибутов (метка 3 на рис.5).

Запись: 17128 из 17128 Представление: Строки Формат: RUSmarc Схема: F Элементы: F

03560nam 2200265 450
001 RU/ВИНИТИ/IN
005 2014120100000.0
100 \$a 20141201d2013 |||y0rusy0179
101 0 \$a rus
102 \$a RU
200 1 \$a Модель самоорганизации в агентных системах с передачей сообщений \$f С.Г. Ломакин
320 \$a 6
330 \$a Доклад посвящен исследованию динамики распространения новой информации (мемов) в социуме (на множестве взаимодействующих агентов - автоматов) на основе модели, рассматривающей распространение мнений (мемов) в графе социальных взаимодействий. Вершины графа представляют агенты (индивиды), каждый имеет список мнений (мемов). В каждый момент времени случайно выбранный индивид озвучивает мнение, случайно выбранное из собственного списка. Слушатели в радиусе досягаемости, сравнивают мнение со своими списками. Если озвученное мнение совпадает с одним из мнений, они с ним соглашаются и стирают остальные. Если нет, дописывают высказанное мнение в список, расширяя свой кругозор. В докладе показано, как преобладающее мнение большинства может быть полностью изменено небольшой частью беспорядочно распределенных агентов, которые распространяют свое мнение и неуязвимы к влиянию других мнений. Также показывается существование переломного момента, после которого сообщество принимает продвигаемую идею. Интересным является факт "пятнистости": при возникновении новой идеи - она распространяется неравномерно на множестве агентов, а агенты объединяются в группы 012001#a5 Международная конференция "Системный анализ и информационные технологии" (САИТ-2013), Красноярск, 19-25 сент., 2013#bT. 1#v978-5-9904056-5-3#C. 243.

Запись: 17128 из 17128 Представление: ISO2709 Формат: RUSmarc Схема: F Элементы: F

03560nam 2200265 450
001# \$a 978-5-9904056-5-3#C. 243.
005 \$a 20141201d2013 |||y0rusy0179 0 agus aRU1 aМодель самоорганизации в агентных системах с передачей сообщений\$fС.Г. Ломакин а6
100 \$a 20141201d2013 |||y0rusy0179 0 agus aRU1 aМодель самоорганизации в агентных системах с передачей сообщений\$fС.Г. Ломакин а6
101 \$1 200 \$a выбранное из собственного списка. Слушатели в радиусе досягаемости, сравнивают мнение со своими списками. Если озвученное мнение совпадает с одним из мнений, они с ним соглашаются и стирают остальные. Если нет, дописывают высказанное мнение в список, расширяя свой кругозор. В докладе показано, как преобладающее мнение большинства может быть полностью изменено небольшой частью беспорядочно распределенных агентов, которые распространяют свое мнение и неуязвимы к влиянию других мнений. Также показывается существование переломного момента, после которого сообщество принимает продвигаемую идею. Интересным является факт "пятнистости": при возникновении новой идеи - она распространяется неравномерно на множестве агентов, а агенты объединяются в группы 012001#a5 Международная конференция "Системный анализ и информационные технологии" (САИТ-2013), Красноярск, 19-25 сент., 2013#bT. 1#v978-5-9904056-5-3#C. 243.
102 \$1 200 \$a выбранное из собственного списка. Слушатели в радиусе досягаемости, сравнивают мнение со своими списками. Если озвученное мнение совпадает с одним из мнений, они с ним соглашаются и стирают остальные. Если нет, дописывают высказанное мнение в список, расширяя свой кругозор. В докладе показано, как преобладающее мнение большинства может быть полностью изменено небольшой частью беспорядочно распределенных агентов, которые распространяют свое мнение и неуязвимы к влиянию других мнений. Также показывается существование переломного момента, после которого сообщество принимает продвигаемую идею. Интересным является факт "пятнистости": при возникновении новой идеи - она распространяется неравномерно на множестве агентов, а агенты объединяются в группы 012001#a5 Международная конференция "Системный анализ и информационные технологии" (САИТ-2013), Красноярск, 19-25 сент., 2013#bT. 1#v978-5-9904056-5-3#C. 243.
200 \$a Модель самоорганизации в агентных системах с передачей сообщений
200 \$a Модель самоорганизации в агентных системах с передачей сообщений
200 \$a Модель самоорганизации в агентных системах с передачей сообщений
200 \$a Модель самоорганизации в агентных системах с передачей сообщений
320 \$a 6
330 \$a Доклад посвящен исследованию динамики распространения новой информации (мемов) в социуме (на множестве взаимодействующих агентов - автоматов) на основе модели, рассматривающей распространение мнений (мемов) в графе социальных взаимодействий. Вершины графа представляют агенты (индивиды), каждый имеет список мнений (мемов). В каждый момент времени случайно выбранный индивид озвучивает мнение, случайно выбранное из собственного списка. Слушатели в радиусе досягаемости, сравнивают мнение со своими списками. Если озвученное мнение совпадает с одним из мнений, они с ним соглашаются и стирают остальные. Если нет, дописывают высказанное мнение в список, расширяя свой кругозор. В докладе показано, как преобладающее мнение большинства может быть полностью изменено небольшой частью беспорядочно распределенных агентов, которые распространяют свое мнение и неуязвимы к влиянию других мнений. Также показывается существование переломного момента, после которого сообщество принимает продвигаемую идею. Интересным является факт "пятнистости": при возникновении новой идеи - она распространяется неравномерно на множестве агентов, а агенты объединяются в группы 012001#a5 Международная конференция "Системный анализ и информационные технологии" (САИТ-2013), Красноярск, 19-25 сент., 2013#bT. 1#v978-5-9904056-5-3#C. 243.

Запись: 17128 из 17128 Представление: Источник Формат: RUSmarc Схема: F Элементы: F

<record xmlns="http://www.loc.gov/MARC21/slim">
<leader>03560nam a2200265 450 </leader>
<controlfield tag="001">RU/ВИНИТИ/IN</controlfield>
<controlfield tag="005">2014120100000.0</controlfield>
<datafield tag="100" ind1="1" ind2="1">
<subfield code="a">20141201d2013 |||y0rusy0179 </subfield>
</datafield>
<datafield tag="101" ind1="0" ind2="1">
<subfield code="a">rus</subfield>
</datafield>
<datafield tag="102" ind1="0" ind2="1">
<subfield code="a">RU</subfield>
</datafield>
<datafield tag="200" ind1="1" ind2="1">
<subfield code="a">Модель самоорганизации в агентных системах с передачей сообщений</subfield>
<subfield code="f">С.Г. Ломакин</subfield>
</datafield>
<datafield tag="320" ind1="1" ind2="1">

Запись: 17128 из 17128 Представление: Обычное Формат: XML Схема: F Элементы: F

Коды ГРНТИ: 20.53.23; УДК: 621.39
Авторы: Ломакин С.Г., Федотов А.М.
Заглавие: Модель самоорганизации в агентных системах с передачей сообщений
Реферат: Доклад посвящен исследованию динамики распространения новой информации (мемов) в социуме (на множестве взаимодействующих агентов - автоматов) на основе модели, рассматривающей распространение мнений (мемов) в графе социальных взаимодействий. Вершины графа представляют агенты (индивиды), каждый имеет список мнений (мемов). В каждый момент времени случайно выбранный индивид озвучивает мнение, случайно выбранное из собственного списка. Слушатели в радиусе досягаемости, сравнивают мнение со своими списками. Если озвученное мнение совпадает с одним из мнений, они с ним соглашаются и стирают остальные. Если нет, дописывают высказанное мнение в список, расширяя свой кругозор. В докладе показано, как преобладающее мнение большинства может быть полностью изменено небольшой частью беспорядочно распределенных агентов, которые распространяют свое мнение и неуязвимы к влиянию других мнений. Также показывается существование переломного момента, после которого сообщество принимает продвигаемую идею. Интересным является факт "пятнистости": при возникновении новой идеи - она распространяется неравномерно на множестве агентов, а агенты объединяются в группы
Ключевые слова: ПЕРЕДАЧА ИНФОРМАЦИИ; РАСПРОСТРАНЕНИЕ ИНФОРМАЦИИ; АГЕНТНЫЕ СИСТЕМЫ; МОДЕЛЬ САМООРГАНИЗАЦИИ; СОЦИАЛЬНЫЕ ВЗАИМОДЕЙСТВИЯ; РАСПРОСТРАНЕНИЕ МНЕНИЙ.
Включено в: 5 Международная конференция "Системный анализ и информационные технологии" (САИТ-2013), Красноярск, 19-25 сент., 2013. - 2013. - Красноярск. - Т. 1. - С. 243-247. - ISBN 978-5-9904056-5-3.

Рис. 7. Различные представления записи: Строки, ISO2709, Источник, Обычное. Представление "Обычное" показывает запись в "дружественной" разметке.

Любое изменение пользователем соответствующего экранного элемента генерирует событие, AJAX-обработчик которого приводит в соответствие новую комбинацию элементов управления с Explain-картой информационного ресурса, хранящейся на WEB-сервере.

Здесь следует отметить, что при работе с множеством информационных источников отображаемый список параметров для каждого элемента управления может соответствовать как минимум двум политикам:

- **Жесткая политика** - список отображаемых параметров формируется как результат логической операции AND, примененной к соответствующим спискам параметров каждого информационного источника.
- **Мягкая политика** - список отображаемых параметров формируется как результат логической операции OR, примененной к соответствующим спискам параметров каждого информационного источника.

При этом использование мягкой политики может приводить к штатным ошибкам для некоторых информационных источников. Тем не менее, подсистема ZooSPACE-W допускает настройку на любую из указанных политик.

При формировании элементов управления для просмотра извлекаемой информации учитывается для каждого информационного ресурса (см. рис. 6)

- Список допустимых форматов представления извлекаемых записей (метка 1 на рис.6)
- Список допустимых схем данных применительно к выбранному формату представления (метка 2 на рис.6)
- Список допустимых форм отображения данных в выбранных форматах и схемах (метка 2 на рис.6)
- Для Z39.50 дополнительно список допустимых элементов данных для выбранной схемы данных (метка 2 на рис.6)

Различные формы отображения записи приведены на рис. 7.

Следует заметить, что использование описанной технологии для построения адаптивных пользовательских интерфейсов в ZooSPACE для работы с гетерогенными источниками данных требует поддержки сервисов Explain для каждого сервера подсистемы ZooSPACE-Z. И если для "родных" серверов ZooPARK-ZS это требование выполняется автоматически, то при интеграции в облако ZooSPACE "чужих" серверов Z39.50 или SRW/SRU (перенаправление запросов специальными провайдерами данных серверов ZooPARK-ZS на другие серверы) требуется или поддержка сервисов Explain этими другими серверами, или ручная генерация информации Explain для этих серверов. Последнее может быть реализовано в подсистеме ZooSPACE-L (LDAP).

В заключение следует заметить, что описанные выше технологии могут быть полезны разработчикам приложений для работы с разнородными информационными ресурсами. Что касается платформы ZooSPACE и всех ее подсистем, то работы по их совершенствованию продолжаются.

Литература

- [1] Жижимов О.Л., Федотов А.М., Шокин Ю.И. Технологическая платформа массовой интеграции гетерогенных данных // Вестник НГУ. Сер.: Информационные технологии. - 2013. - Т.11. - № 1. - С.24-41. - ISSN 1818-7900.
- [2] Жижимов О. Л., Лобыкин А. А., Турчановский И. Ю., Паньшин А. А., Чудинов С. А. Автоматизированная система сбора статистической информации о событиях в распределенной информационной системе //

Вестник НГУ. Сер.: Информационные технологии. - 2013. - Т.11. - № 1. - С.42-52. - ISSN 1818-7900.

- [3] Жижимов О.Л., Федотов А.М., Шокин Ю.И. Платформа ZooSPACE - организация доступа к разнородным распределенным ресурсам // Электронные библиотеки: российский научный электронный журнал. - 2014. - Т.17. - № 2. - ISSN 1562-5419.
- [4] Dahl M., Banerjee K., Spalti M. Digital Libraries: Integration Content and System. - 2015 (2006). - Chandos Pub 01/06/2015. - ISBN 1843341557.
- [5] Жижимов О.Л. О методах сравнения информационных систем, их классификации и специальных сервисах [Электронный ресурс] // Распределенные информационные и вычислительные ресурсы (DICR-2010): материалы XIII Российской конференции с международным участием (электронное издание). - 2010. - Новосибирск: Институт вычислительных технологий Сибирского отделения РАН. - Гос. регистр. №: 0321100051.
- [6] ANSI/NISO Z39.50-2003. Information Retrieval (Z39.50): Application Service Definition and Protocol Specification. - NISO Press, Bethesda, Maryland, U.S.A. — Nov 2002. — ISSN: 1041-5653. — ISBN: 1-880124-55-6.
- [7] SRU - Explain Operation / The Library of Congress, 6.09.2013 - <http://www.loc.gov/standards/sru/explain/>
- [8] ZeeRex: The Explainable "Explain" Service - <http://zeerex.z3950.org>
- [9] SRU- Search/Retrieve via URL / The Library of Congress, 26.12.2013 - <http://www.loc.gov/standards/sru/>
- [10] Zebra - User's Guide and Reference / Index Data, 1994-2009 - <http://www.indexdata.com/zebra/doc/zebra.pdf>

Explain Services on ZooSPACE Platform and Adaptive User Interfaces

Oleg L. Zhizhimov

The technology implementation services Explain on the platform ZooSPACE, which provides access to distributed heterogeneous information resources in accordance with the protocols SRW/SRU, including Z39.50 through the graphical user WEB-interface are considered. We discuss the details of implementing user interfaces for information retrieval in a heterogeneous information resources that adapts to the structure and functionality of information sources on the example subsystem ZooSPACE-W.