Интеллектуализация многофункциональной электронной библиотеки: концепция, методология, технология

С.Х.Ляпин, к.ф.н, доц.

ООО «Константа», НП «Центроконцепт»
lyapin@atknet.ru
Архангельский областной центр повышения квалификации специалистов культуры (АОЦПК),
http://www.arkhadm.gov.ru/culture;
cpk@atnet.ru

Международная научная конференция «Интеллектуальные информационные технологии в библиотечном деле», 11-12 октября 2005 г., Киев, Национальная библиотека Украины им. В.И.Вернадского.

Содержание

- Введение. От предоставления ресурсов – к многофункциональности и интеллектуальности.
- Многоцелевая информационная среда T-System, ее общая платформа T-Libra и специализированные блоки T-Archive, T-Media, T-Research, T-Education.
- Проблема интеллектуализации: ориентировочные этапы (от «объектов» к «агентам»).
- Проблема интеллектуализации: концептуальный, методологический и технологический аспекты.
- Методологические ориентиры для онтологии-I.
- Технологические ориентиры для онтологии-I.
- Заключение. О внедрении разработок и воспитании пользователя.
- Литература.

Введение. От предоставления ресурсов - к многофункциональности и интеллектуальности (1)

- **Электронные библиотеки, виртуальные музеи, электронные архивы и т.п., начав с предоставления доступа к ресурсам, активно развиваются в направлении продвинутых поисковых сервисов по многофункциональной обработке и презентации как текстовой, так и нетекстовой информации.**
- **Примеры**:
  - **американские проекты «Alexandria Digital Earth Prototype»** (ADEPT) **[1]** и «**Digital Library for Earth System Education**» (DLESE) **[2]**, ориентированные на комплексную поддержку обучения и исследований в сфере наук о Земле;
  - **американо-германский проект «Archimedes»** **[3]**, направленный на создание интерактивной среды по истории механики;
  - **Электронный архив академика А.П.Ершова** (СО РАН) **[4]**;
  - Распределенная информационная система «**Виртуальный музей науки и техники СО РАН**» **[5]**;
  - **Информационная система T-Libra** (ООО «Константа», г. Архангельск), вначале разработанная для создания электронных библиотек с гибким тематизируемым полнотекстовым поиском **[6]**.

Введение. От предоставления ресурсов – к многофункциональности и интеллектуальности (2)

- **Обеспечение многофункциональности информационной среды тесно связано с ее *интеллектуализацией* – как с точки зрения общей архитектуры и организации первичных ресурсов, так и с точки зрения поисковых и презентационных сервисов.**
- **Интеллектуализация становится критически важным условием рыночной выживаемости многофункциональных интегративных информационных систем в средне- и долгосрочной перспективе**.

Тема доклада

- **Развитие ИС T-Libra в направлении многоцелевой информационной среды T-System, сочетающей в себе функции электронной библиотеки, электронного архива, виртуального музея, исследовательской лаборатории, образовательного сервера**, путем
  - **комплексной интеллектуализации информационной среды на основе моделирования «логики» пользовательского запроса**
  - **и поэтапного переноса интеллекта от «человека-пользователя» к создаваемой многоагентной среде (MAC) и «интеллектуальной организации».**

Введение. От предоставления ресурсов – к многофункциональности к интеллектуальности (3)

Социум, естественный и искусственный интеллект

- «Библиотека», «Архив», «Экспозиция/Музей», «Исследование», «Обучение» - обобщенные социальные и интеллектуальные функции, представляющие многоплановую и социально-значимую часть пространства человеческой деятельности (в тесно связанных между собой сферах культуры, науки, образования).
- «Интеллектуализация» соответствующих интегрированных информационных сред – это одновременно:
  - условие их востребованности обществом,
  - полигон для апробации теоретических положений и прикладных разработок как в сфере естественного, так и искусственного интеллекта.

Содержание

- Введение. От предоставления ресурсов – к многофункциональности и интеллектуальности.
- Многоцелевая информационная среда T-System, ее общая платформа T-Libra и специализированные блоки T-Archive, T-Media, T-Research, T-Education.
- Проблема интеллектуализации: концептуальный, методологический и технологический аспекты.
- Этапы интеллектуализации: от «объектов» к «агентам». Семейство онтологий для T-System.
- Методологические ориентиры для онтологии-I.
- Технологические ориентиры для онтологии-I.
- Заключение. О внедрении разработок и воспитании пользователя.
- Литература.

Многоцелевая информационная среда T-System: основные функциональные блоки (1)

T-System

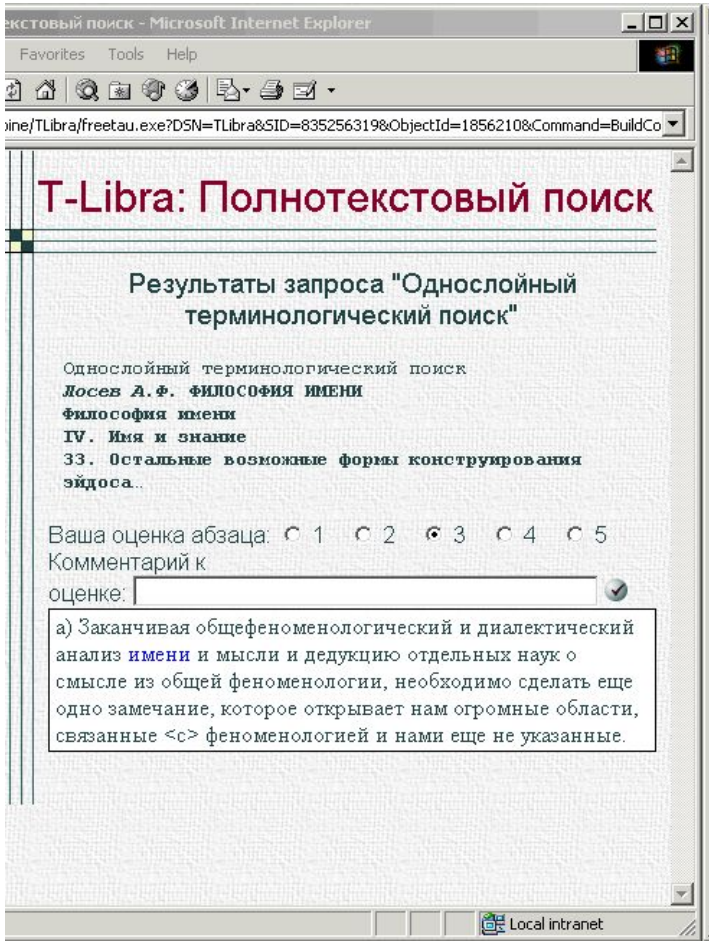
- **Общая платформа**: **электронная библиотека T-Libra** с интегрированным каталогом, мультимодальным депозитарием и гибким тематизируемым полнотекстовым поиском по полнотекстовым ресурсам, существующим в формате объектно-реляционных баз данных
- **Электронный архив T-Archive** (+ возможность создания тематических коллекций документов с навигацией по ним)
- **Информационная среда T-Media** (+ возможность создания интерактивных мультимодальных тематических экспозиций для виртуальных музеев, преподавания учебных дисциплин и т.п.)
- **Информационная среда T-Research** (+ возможность создания виртуальных исследовательских лабораторий и установок)
- **Информационная среда T-Education** (+ образовательный сервер для поддержки автоматизированного, в том числе дистантного обучения)

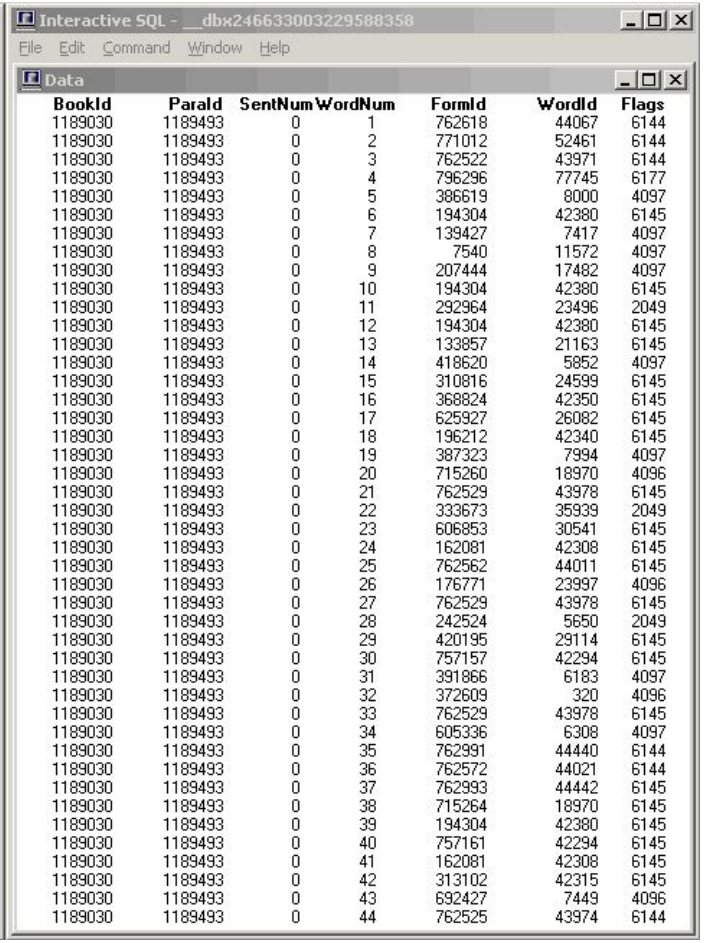
Многоцелевая информационная среда T-System: основные функциональные блоки (2)

Информационная система T-Libra

- **Для создания электронных библиотек с пополняемым (многоязычным) словарем словоформ, интегрированным каталогом, мультимодальным депозитарием и гибким тематизируемым поиском по полнотекстовым ресурсам, существующих в формате объектно-реляционных баз данных (прототип – 2000 г, первое внедрение – 2002 г.).**
- Варианты с **двух- и трехзвенной клиент-серверной Интернет-архитектурой** (без сервера приложений и с сервером приложений).
  - **Внутренняя логика** с учетом специфики СУБД (есть реализации для MS SQL Server и Sybase ASA, версии 4.x – 5.x)
  - **Внешняя логика** (начиная с версии 6.x), обеспечивающая реальную многоплатформенность в плане независимости от используемой СУБД
- **Литература по развитию ИС T-Libra [6].**
- **Поисковая машина Яндекс (Yandex)** по запросу “T-Libra” выдает более десятка ссылок (в основном, на тексты тезисов и статей на крупных конференциях – **RCDL, EVA-Moscow, Научный сервис в сети Интернет и т.д.**)

Многоцелевая информационная среда
T-System: основные функциональные блоки (3)
T-Libra: два стандартных формата представления текста в Интернет-архитектуре (html-файл и реляционная БД)





BookId	ParaId	SentNum	WordNum	FormId	WordId	Flags
1189030	1189493	0	1	762618	44067	6144
1189030	1189493	0	2	771012	52461	6144
1189030	1189493	0	3	762522	43971	6144
1189030	1189493	0	4	796296	77745	6177
1189030	1189493	0	5	386619	8000	4097
1189030	1189493	0	6	194304	42380	6145
1189030	1189493	0	7	139427	7417	4097
1189030	1189493	0	8	7540	11572	4097
1189030	1189493	0	9	207444	17482	4097
1189030	1189493	0	10	194304	42380	6145
1189030	1189493	0	11	292964	23496	2049
1189030	1189493	0	12	194304	42380	6145
1189030	1189493	0	13	133857	21163	6145
1189030	1189493	0	14	418620	5852	4097
1189030	1189493	0	15	310816	24599	6145
1189030	1189493	0	16	368824	42350	6145
1189030	1189493	0	17	625927	26082	6145
1189030	1189493	0	18	196212	42340	6145
1189030	1189493	0	19	387323	7994	4097
1189030	1189493	0	20	715260	18970	4096
1189030	1189493	0	21	762529	43978	6145
1189030	1189493	0	22	333673	35939	2049
1189030	1189493	0	23	606853	30541	6145
1189030	1189493	0	24	162081	42308	6145
1189030	1189493	0	25	762562	44011	6145
1189030	1189493	0	26	176771	23997	4096
1189030	1189493	0	27	762529	43978	6145
1189030	1189493	0	28	242524	5650	2049
1189030	1189493	0	29	420195	29114	6145
1189030	1189493	0	30	757157	42294	6145
1189030	1189493	0	31	391866	6183	4097
1189030	1189493	0	32	372609	320	4096
1189030	1189493	0	33	762529	43978	6145
1189030	1189493	0	34	605336	6308	4097
1189030	1189493	0	35	762991	44440	6144
1189030	1189493	0	36	762572	44021	6144
1189030	1189493	0	37	762993	44442	6145
1189030	1189493	0	38	715264	18970	6145
1189030	1189493	0	39	194304	42380	6145
1189030	1189493	0	40	757161	42294	6145
1189030	1189493	0	41	162081	42308	6145
1189030	1189493	0	42	313102	42315	6145
1189030	1189493	0	43	692427	7449	4096
1189030	1189493	0	44	762525	43974	6144

Многоцелевая информационная среда T-System: основные функциональные блоки (4)

Гибкий поиск в T-System
(T-Libra, примеры запросов)

1. **Экспликация тематических структур текста в составе авторского абзаца** (параметризируемый многослойный тематический поиск).
2. **Конкорданс** (произвольное, грамматически неизменяемое словосочетание в составе предложения / абзаца).
3. **Терминограммы произведения, абсолютные и относительные** (построение частотно-ранжированных списков имен существительных по выбранному произведению или группе произведений, приведенных к нормальной форме).

Многоцелевая информационная среда T-System: основные функциональные блоки (5)

Гибкий поиск в T-System (T-Media, T-Research – примеры запросов)

- **4. Построение мультимодальных экспозиций (ручное и автоматизированное) с активным использованием полнотекстового поиска. T-Media**

- **5. Генерация грамматической парадигматики слов русского языка (правильные и неправильные парадигмы). Изучение выразительного потенциала языка (в плане словоизменения). T-Research**

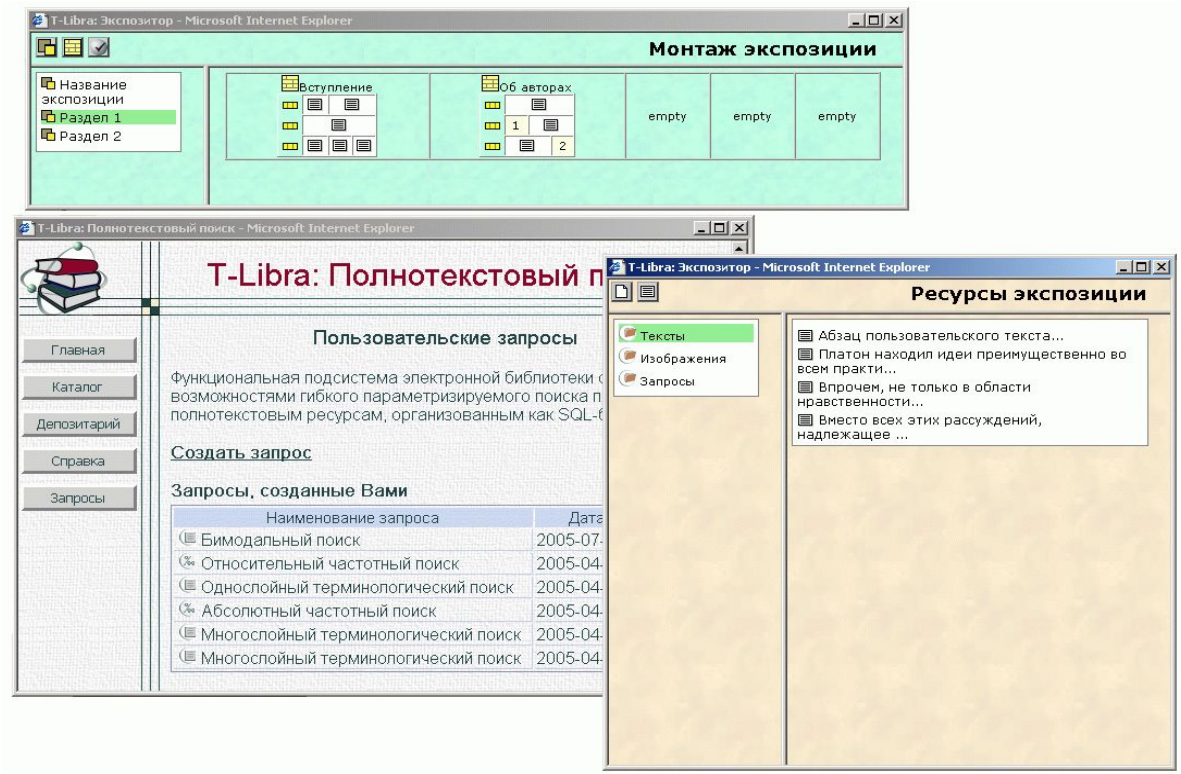
- **6. Статическое и динамическое моделирование семантических структур текста. Изучение культурных концептов, существующих в редуцированной текстовой форме. T-Research**

Многоцелевая информационная среда T-System: основные функциональные блоки (6)

Гибкий поиск в T-System (T-Media, модуль Expositor)

- Создание тематических мультимодальных экспозиций для различных целей (виртуальный музей, образовательная среда и т. д.). T-Media
- При запуске модуля Expositor открываются три окна для ручного создания экспозиции:
  - окно «Полнотекстовый поиск» (функциональный блок пользовательского раздела T-Libra),
  - две опции собственно модуля Expositor, окно «Ресурсы экспозиции» и окно «Монтаж экспозиции».
- Тематическая экспозиция может быть составлена также в автоматизированном режиме (путем выполнения каскадного запроса, в том числе концепт-ориентированного).
- Пример реальной работы с модулем Expositor (T-Media)

Многоцелевая информационная среда T-System: основные функциональные блоки (7)



- **T-Media = T-Libra + Expositor (модуль создания интерактивных мультимодальных тематических экспозиций, взаимодействующих с электронной библиотекой: для виртуальных музеев, преподавания и т.д.)**

Многоцелевая информационная среда T-System: основные функциональные блоки (8)

Гибкий поиск в T-System
(T-Research, модуль Virtula для создания виртуальной лаборатории)
Пример 1:

- **Генерация грамматической парадигмы слов русского языка (T-Research).**
- **Программный комплекс ParaGen для любого слова русского языка, введенного исследователем:**
  - **а) выдвигает гипотезу о принадлежности этого слова к той или иной части речи;**
  - **б) восстанавливает в рамках этой гипотезы базисную форму слова;**
  - **в) генерирует формально возможный набор словоформ для данного слова, т.е. его грамматическую парадигму;**
  - **г) выбирает из всех сгенерированных парадигм правильную;**
  - **д) показывает «правильную» парадигму в контексте всех «неправильных».**
- **Может использоваться для автоматизированного пополнения электронных словарей словоформ; для анализа процесса словообразования; для анализа выразительного потенциала языковых форм (в плане словообразования); для задач компаративной лингвистики.**
- **Пример реальной работы программного комплекса ParaGen v.2.0**

Многоцелевая информационная среда T-System: основные функциональные блоки (9)

T-Libra: Виртуальная лаборатория - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Address http://pine/TLibraScript/freetau.exe?DSN=TLibra&SID=1813381119&ObjectId=1012&Command=BuildParagenPlant&JobId=333&HypNum=1&Hyp

T-Libra: Виртуальная лаборатория

Закрыть

ВЭУ Анализ и моделирование словоизменения

ВЭУ для любого слова русского языка: • выдвигает гипотезы о принадлежности этого слова к той или иной части речи и определяет соответствующие грамматические признаки; • восстанавливает в рамках этих гипотез базисную форму слова; • генерирует формально возможный набор словоформ для данного слова (грамматическую парадигму).

Исходное слово: философия Создать

Список построенных гипотез

№	Слово	Часть речи	Грамматические признаки
1	**философия**	Имя существительное	Женский род, Первое склонение, Именительный падеж, Число единственное
2	философий	Имя существительное	Первое склонение, Именительный падеж, Мужской род, Число единственное
3	философие	Имя существительное	Первое склонение, Именительный падеж, Число единственное
4	философие	Имя существительное	Первое склонение, Именительный падеж, Число единственное
5	философия	Имя существительное	Второе склонение, Именительный падеж, Число единственное

Парадигма №1 (JobId=333, HypId=1)

Часть речи:	Имя существительное
Генерация парадигмы проведена процедурой:	GenerateFilos

Падеж	Ед.ч.	Мн.ч.
Им.	философия	философии
Род.	философии	философий
Дат.	философии	философиям
Вин.	философия	философии
Тв.	философией	философиями
Пр.	философии	философиях

Done Local intranet

Многоцелевая информационная среда T-System: основные функциональные блоки (10)

Гибкий поиск в T-System (T-Research, модуль Virtula для создания виртуальной лаборатории)
Пример 2:

- **Моделирование культурных концептов (существующих в редуцированной текстовой форме). T-Research**
  - **а) Статическое. Содержательные параметры поиска (тезаурус II рода) задаются пользователем в специальной форме запроса**
  - **б) Динамическое. Содержательные параметры поиска автоматически формируются в ходе выполнения запроса (задаются только технические параметры)**
- **Пример реального поиска (T-Research)**

Содержание

- Введение. От предоставления ресурсов – к многофункциональности и интеллектуальности.
- Многоцелевая информационная среда T-System, ее общая платформа T-Libra и специализированные блоки T-Archive, T-Media, T-Research, T-Education.
- Проблема интеллектуализации: ориентировочные этапы (от «объектов» к «агентам»).
- Проблема интеллектуализации: концептуальный, методологический и технологический аспекты.
- Методологические ориентиры для онтологии-I.
- Технологические ориентиры для онтологии-I.
- Заключение. О внедрении разработок и воспитании пользователя.
- Литература.

Проблема интеллектуализации T-System: основные этапы (1)

Четыре этапа интеллектуализации T-System (сообразно концептуальным моделям «пользователя»)

- **1. Пользователь = человек-индивид с прямым «ручным» управлением интеллектуальным поиском, формирующий функциональные структуры запроса и выбирающий формы презентации его результата.**
- **2. Пользователь = человек-индивид + отдельные специализированные программные агенты (не составляющие самоорганизующуюся среду); частичная передача управления поиском (сервисами) агентам.**
- **3. Пользователь = человек-индивид + многоагентная система (МАС), составляющая самоорганизующуюся среду; в основном управление передается агентной среде.**
- **4. Пользователь = «интеллектуальная организация»: все управление осуществляется самоорганизующейся многоагентной системой (МАС); пользователь («индивид» или «виртуальный коллектив») – один из ее функциональных фрагментов.**
  - **[7]** *В.Б.Тарасов*. От многоагентных систем к интеллектуальным организациям: философия, психология, информатика. – М.: Эдиториал УРСС, 2002. -352с

Проблема интеллектуализации T-System: основные этапы (2)

1-й этап: основные характеристики

- **Пользователь = «человек-индивид».**
- **Прямое управление поиском путем «ручного» формирования функциональных структур интеллектуального запроса и выбора форм презентации его результата.**
- **Тип интеллектуализации: классическая знаниецентрическая (когнитивная) модель индивидуального интеллекта, реализованная в онтологии-I (моделируются «знания»).**
- **Разработки ведутся в рамках объектно-ориентированного программирования (ООП) с элементами параллелизма (параллельная обработка централизованных данных).**
- **Централизованное размещение данных (сервер базы данных с удаленным доступом).**

Проблема интеллектуализации T-System: основные этапы (3)

2-й этап: основные характеристики

- **Пользователь = «человек-индивид» + отдельные специализированные программные агенты (не составляющие самоорганизующуюся среду).**
- **Частичная передача управления интеллектуальным поиском (отдельными интеллектуальными сервисами) агентам.**
- **Тип интеллектуализации: неклассическая интенциональная модель интеллекта, реализованная в онтологии-II (моделируются не только «знания», но и «цели», «ценности», «намерения» и т.д.).**
- **Разработки в рамках сочетания объектно-ориентированного программирования (ООП) и агентно-ориентированного программирования (АОП).**
- **Элементы параллелизма (параллельная обработка централизованных данных).**
- **Централизованное размещение данных (сервер базы данных с удаленным доступом).**

Проблема интеллектуализации T-System: основные этапы (4)

3-й этап: основные характеристики

- **Пользователь = «человек-индивид» + «многоагентная система» (МАС), составляющая самоорганизующуюся среду;**
- **Функции управления среднего («знания», «цели», «ценности» и т. д.) и нижнего («данные») уровней передаются многоагентной среде.**
- **Тип интеллектуализации: неклассическая кибернетическая модель индивидуального интеллекта (нелинейность, нечеткая логика), реализованная в онтологии-III.**
- **Разработки в рамках агентно-ориентированного программирования (АОП) с использованием распределенных параллельных вычислений (параллельная обработка данных, параллельное выполнение запроса).**
- **Распределенное размещение и обработка данных; активное использование GRID-технологий (или им подобных).**

Проблема интеллектуализации T-System: основные этапы (5)

4-й этап: основные характеристики

- Пользователь = «интеллектуальная организация».
- Функции управления всех уровней передаются самоорганизующейся многоагентной системе (МАС) и в целом «интеллектуальной организации».
- Обычный пользователь («человек-индивид» или «виртуальный коллектив») – один из функциональных фрагментов «интеллектуальной организации».
- Тип интеллектуализации = неклассическая кибернетическая модель коллективного интеллекта = «интеллектуальная организация», реализованная в онтологии-IV.
- Разработки в рамках распределенного искусственного интеллекта (РИИ) и распределенных параллельных вычислений.
- Сетевая организация и обработка данных, интероперабельность («прозрачность» гетерогенной глобальной сети для выполнения произвольных запросов). Интернет + GRID + …

Содержание

- Введение. От предоставления ресурсов – к многофункциональности и интеллектуальности.
- Многоцелевая информационная среда T-System, ее общая платформа T-Libra и специализированные блоки T-Archive, T-Media, T-Research, T-Education.
- Проблема интеллектуализации: ориентировочные этапы (от «объектов» к «агентам»).
- Проблема интеллектуализации: концептуальный, методологический и технологический аспекты.
- Методологические ориентиры для онтологии-I.
- Технологические ориентиры для онтологии-I.
- Заключение. О внедрении разработок и воспитании пользователя.
- Литература.

Проблема интеллектуализации T-System: три аспекта (1)

Три аспекта проблемы

- **Концептуальный: рекурсивно-конструктивный подход, оформленный как программа поэтапного и покомпонентного переноса «интеллекта» от человека/пользователя к информационной среде (ИС) в условиях ее реального функционирования, с поэтапной и покомпонентной оценкой адекватности и релевантности такого переноса.**
- **Методологический: методологические ориентиры (модели, эвристики) для интеллектуализации ИС, прежде всего для проектирования и конкретизации ее онтологической модели**
- **Технологический: выбор архитектуры информационной среды, организации ресурсов, подсистем поиска и презентации его результатов.**

- **Ключевой момент всей проблематики: моделирование «логики» пользовательского запроса в многоцелевой среде.**

Проблема интеллектуализации T-System: концептуальный аспект (2)

Концептуальный аспект проблемы

- В конфигурации **[Человек (пользователь) + Информационная система с ресурсами и поиском]** присутствует «интеллект» как минимум в виде **человеческого интеллекта**, способного ставить цель поиска, реализовывать ее в поисковом запросе и оценивать его результаты.
- **Задача**: **неявные характеристики интеллекта, присущие левой части конфигурации («человек»), поэтапно и покомпонентно превратить в явные характеристики интеллекта самой информационной среды («интеллектуальная организация»).**
- Формирование интеллекта ИС идет под контролем **рекурсивно-конструктивного переноса** функциональных структур интеллекта (от «человека/пользователя» к ИС и обратно), с экспертной оценкой каждой рекурсии по результатам работы реальных поисковых запросов.
- **Для этого нужны**:
  - **интерпретация** различных концепций и моделей человеческого интеллекта (философских, общесистемных, психологических и др.) **в терминах и «логике» информационной среды**,
  - **реинтерпретация (коррекция)** философских, общесистемных психологических и др. концепций человеческого интеллекта в связи с оценкой эффективности работы интеллектуально нагруженной ИС.

Содержание

- Введение. От предоставления ресурсов – к многофункциональности и интеллектуальности.
- Многоцелевая информационная среда T-System, ее общая платформа T-Libra и специализированные блоки T-Archive, T-Media, T-Research, T-Education.
- Проблема интеллектуализации: ориентировочные этапы (от «объектов» к «агентам»).
- Проблема интеллектуализации: концептуальный, методологический и технологический аспекты.
- Методологические ориентиры для онтологии-I.
- Технологические ориентиры для онтологии-I.
- Заключение. О внедрении разработок и воспитании пользователя.
- Литература.

Проблема интеллектуализации T-System: методологический аспект (1)

Методологические ориентиры для онтологии

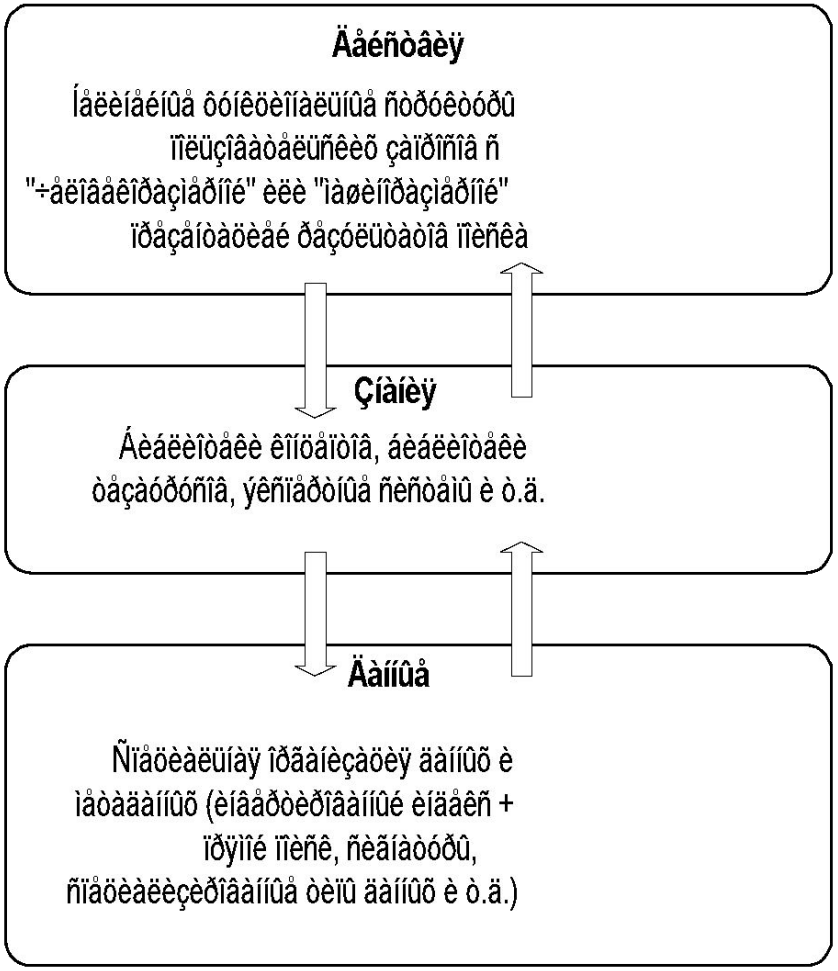
- **Методологические ориентиры (модели, эвристики) из сфер философии, теории систем, кибернетики, психологии, лингвистики и т.д. для информационного моделирования «интеллекта» выбираются не вообще, а под некоторую *онтологическую модель*.**
- **Для поэтапной комплексной интеллектуализации необходимо *семейство онтологий* (по крайней мере, по одной онтологии на каждый этап).**
- **Специфика онтологической модели для 1 этапа интеллектуализации (Онтология-I):**
  - **Обобщенные координаты / уровни модели: *«действия» – «знания» – «данные»*.**
  - **Знаниецентрическая (когнитивная) модель интеллекта информационной среды с элементами деятельностного подхода (необходимо для функциональной интеграции поиска).**

Проблема интеллектуализации T-System: методологический аспект (2)

Три уровня организации информационной среды (для Онтологии-I)

- ***верхний уровень («действия»),*** – управленческий уровень, на нем представлены функциональные структуры пользовательских запросов и специальные средства для репрезентации результатов поиска как в человекоразмерной форме (результат поиска для использования его человеком), так и в машиноразмерной (результат поиска для дальнейшего использования его программным агентом).
- ***средний уровень («знания»),*** – содержательный уровень, на нем представлены различные формы вторичной организации информации: мультимодальные культурные концепты (библиотеки концептов), предметно-ориентированные тезаурусы (библиотеки тезаурусов), встроенные экспертные системы, аналитические и оценочные модули, специализированные базы знаний и т.п.
- ***нижний уровень («данные»),*** – представлены специализированные формы организации первичной информации для обеспечения эффективного интеллектуального поиска (сочетание инвертированного индекса с прямым поиском, сигнатуры, специальные типы данных и метаданных …).

Проблема интеллектуализации T-System: методологический аспект (3).



Онтология-I

- Трехуровневая блок-схема онтологической модели для 1-го этапа разработки «интеллекта» информационной среды T-System
- Основные элементы Онтологии-I необходимы, но недостаточны для Онтологий II, III, IV

Методологические ориентиры для верхнего уровня онтологии-I («действия») (4)

Основной методологический ориентир

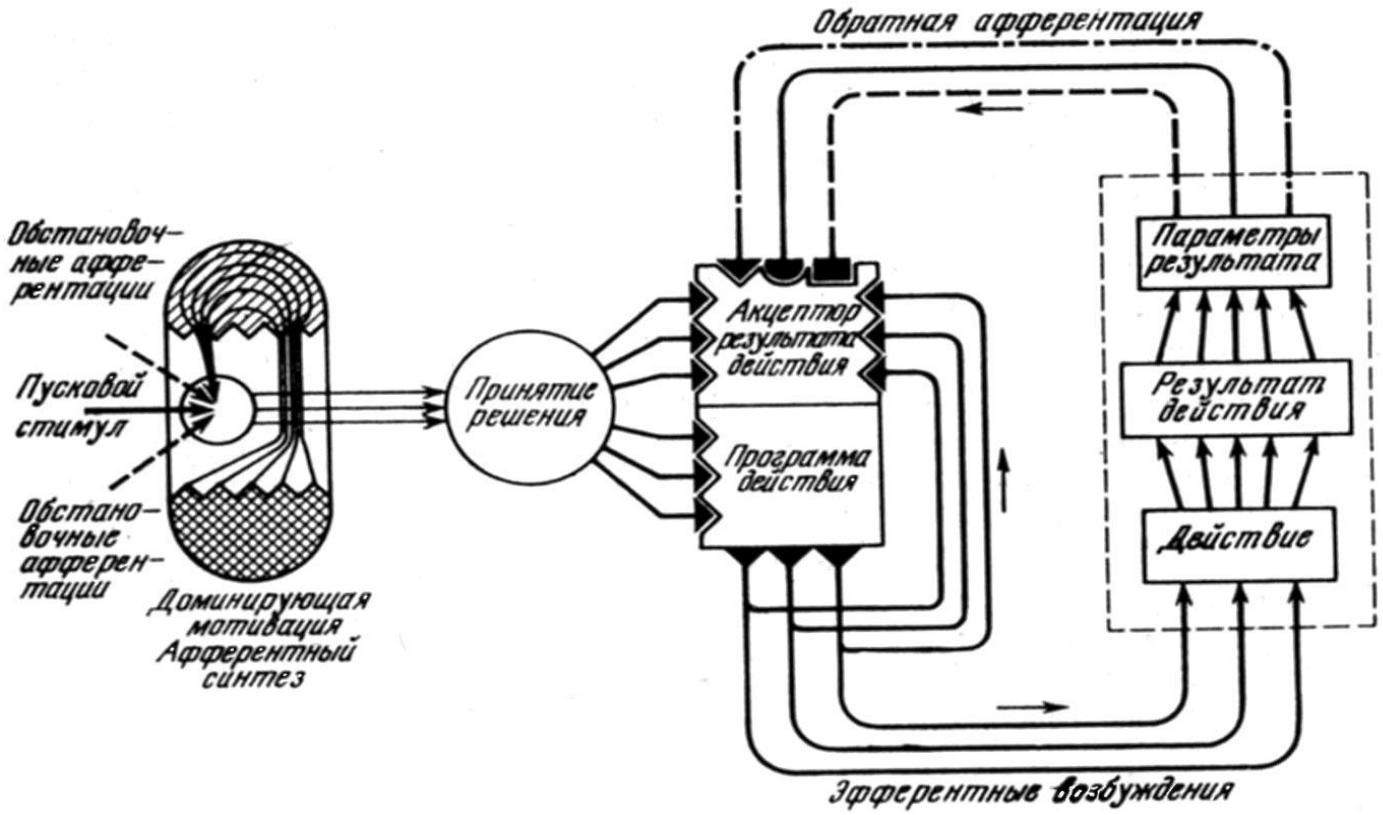
- Основной единицей верхнего уровня онтологии является ***функциональная структура пользовательского запроса***.

- **Примечание**. Эта же единица является интегратором интеллектуального поиска (интеллектуальной информационной среды) в целом.

- Основной методологический ориентир для проектирования верхнего уровня онтологии-I: ***теория функциональной системы П.К.Анохина*** **[8].**

Методологические ориентиры для верхнего уровня онтологии-I («действия») (5)

Общая архитектура функциональной системы по П.К.Анохину



Методологические ориентиры для верхнего уровня онтологии-I («действия») (6)

Задача интерпретации схемы П.К.Анохина

- Ключевые компоненты анохинской функциональной системы:
  - *результат* как доминирующий системообразующий фактор;
  - *афферентный синтез*, предшествующий поведенческому акту;
  - *обратная афферентация* (обратная связь);
  - *аппарат акцептора результата действия*;
  - *нелинейное взаимосодействие* всех этих компонентов в составе целостного поведенческого акта

- должны получить **интерпретацию в терминах информационной среды** (прежде всего - **в терминах пользовательского запроса** к определенным образом организованным первичным ресурсам).

Методологические ориентиры для верхнего уровня онтологии-I («действия») (7)

«Эпистемология» в составе онтологии

- Верхний уровень онтологии-I можно назвать «эпистемологическим»:
  - функциональные системы, формируемые поисковыми запросами, фактически репрезентируют ту или иную *функциональную организацию «субъекта»* (реального пользователя и/или программного агента, формирующих запрос), и, соответственно, *нелинейную функциональную организацию субъект-объектного отношения*.

- В общеметодологическом плане (и тоже со ссылками на теорию П.К.Анохина) это сформулировано нами еще в начале 1980-х гг. в рамках концептуальной модели «совокупного субъекта», или Σ-субъекта (сигма-субъекта) **[9]**

Методологические ориентиры для верхнего уровня онтологии-I («действия») (8)

Дополнительные методологические ориентиры

- **Нелинейная эпистемология*:***
  - **Предметно-редуктивная теория сознания (теория «идеального»), восходящая к К.Марксу (Э.В.Ильенков, М.А. Лифшиц, М.А.Мамардашвили…),**
  - **экзистенциально-феноменологическая традиция, в особенности Э.Гуссерль (ноэма, интенциональная структура сознания) и М.Хайдеггер (Dasein, бытийно-понимающая экзистенция);**
  - **Теория сознания «позднего» М.К.Мамардашвили (рефлексия как актуализатор активных состояний сознания и познания).**
- **Генетическая эпистемология (Ж.Пиаже), эволюционная эпистемология (К.Поппер).**
- **Теория метасистемных переходов и кибернетическая эпистемология (В.Ф.Турчин [10]).**
- **Теории интеллектуального управления, прикладная семиотика (Д. А.Поспелов, В.Г.Редько, …) [11]**

Методологические ориентиры для среднего уровня онтологии-I («знания») (9)

«Знаниецентрическая» модель интеллекта для среднего уровня онтологии-I

- Основной единицей среднего уровня онтологии является та или иная **форма организации знания**.
- Базовые элементарные формы организации знания, адекватные задачам организации гибкого поиска: «**библиотеки концептов**» и «**библиотеки тезаурусов**»
  - ***Библиотека концептов*** состоит из многомерных мультимодальных концептов («ММ-концептов»), - смысловых единиц информации, - и ориентирована на репрезентацию межпредметных и междисциплинарных форм знания, познания, сознания;
  - ***Библиотека тезаурусов*** состоит из специализированных тезаурусов (локальных или соответствующих фрагментов глобальных) и ориентирована на понятийно-терминологическую репрезентацию предметно-дисциплинарных областей знания

Методологические ориентиры для среднего уровня онтологии-I («знания») (10)

Основные ориентиры:

- **Общая теория концептов (*концептология*), семиотика, логический анализ языка, теория информационных тезаурусов.**

- **Психологическое понимание «понятийного концепта», его содержательная структура и интеллектообразующая роль (Л.М.Веккер [12])**

- **Нечеткая логика, мягкие вычисления (Information Granulation и Computing with Words) (Л.Заде [13])**

Методологические ориентиры для среднего уровня онтологии-I («знания») (11)

Основные ориентиры:

теория концептов (и тезаурусов)

- Концепты как многомерные и мультимодальные смысловые единицы информации, существующие в различных редуцированных знаковых (в том числе терминологических) формах (**С.Х.Ляпин [14]**).
- Экспликация и описание концептов как культурно-языковых констант (**Ю.С.Степанов [15]**).
- Концепты как логические пропозиции (пропозициональные функции) языковых явлений (**Н.Д.Арутюнова [16]**).
- Тезаурусы как содержательное отображение предметной области в форме понятийно-терминологических структур (**Ю. А.Шрейдер, …**).

Методологические ориентиры для среднего уровня онтологии-I («знания») (12)

Основные ориентиры:
модель «понятийного концепта» по Л.М.Веккеру

- **Психологический анализ «понятия» как двойного инварианта обратимого межъязыкового перевода (с языка свернутых симультанных образов на язык линейно упорядоченных речевых символов), ведущегося минимум на двух уровнях родо-видовой обобщенности.**

- ***Понятийный концепт* как «центр кристаллизации» других понятийных концептов и тем самым – интеллектообразующая единица.**

- **Операциональная структура понятийного концепта (понимаемого по Веккеру) - основной ориентир для конструирования многомерного мультимодального (культурного) концепта и соответствующих интеллектуальных структур более высокого (надличностного) уровня.**

Методологические ориентиры для среднего уровня онтологии-I («знания») (13)

Основные ориентиры:
Information Granulation и Computing with Words (Л.Заде)

- **Гранулирование информации. Слово естественного языка понимается как «метка» гранулы, а «гранула» – как группа объектов, обозначаемых словом и объединяемых неразличимостью, сходством, близостью или функциональностью.**
- **Организацию знания в виде «концептов» и «тезаурусов» можно понимать как частный случай гранулирования. Соответственно, можно использовать язык описания гранулированного знания для описания концепт-организованного знания или предметно-ориентированных тезаурусов.**
- **Вычисления со словами. Слова рассматриваются как ограничения на *лингвистическую переменную*, и основной компонентой процесса вычисления со словами является распространение ограничений с одних переменных на другие.**
- **Распространение подхода на теоретическую и технологическую разработку «вычислений с концептами» *(Computing with Concepts)*, – операции производятся с концептами как целостными формообразованиями с инкапсулированной логикой поведения в культурном поле. С соответствующими изменениями можно говорить, видимо, и о «вычислениях с тезаурусами» *(Computing with Thesauri)*.**

Методологические ориентиры для среднего уровня онтологии-I («знания») (14)

Дополнительные ориентиры:

- **Лингвистика текста** (экспликация сверхфразовых единиц текста, вертикальных структур текста и т.д.)
- **Постмодернистская текстология** (Ж.Деррида, Р.Барт, П.Рикер, Ж.Делез, Ю.Кристева, …), включая также работы по семиотике Ю.М.Лотмана, В.Н.Топорова, В.А.Успенского
  - **«Текст» (в широком культурологическом смысле слова) как децентрализованный самоорганизующийся открытый культурный феномен с сетевой генерацией значений и множественностью смыслов[17].**
- **Нейроинформатика, нейросетевой подход.**
  - **а) нейросетевая парадигма для организации информационных ресурсов (мультимодальных концептов, локальных тезаурусов и их совокупностей), а нейросетевых алгоритмов для организации поиска (включая распознавание смысла, в том числе смысла текста),**
  - **б) эвристическое моделирование информационной среды в целом с точки зрения *коннекционизма*, – то есть рассмотрение *связей* структуры в качестве хранителей и носителей ее основных свойств [18].**

Методологические ориентиры для нижнего уровня онтологии-I («данные») (15)

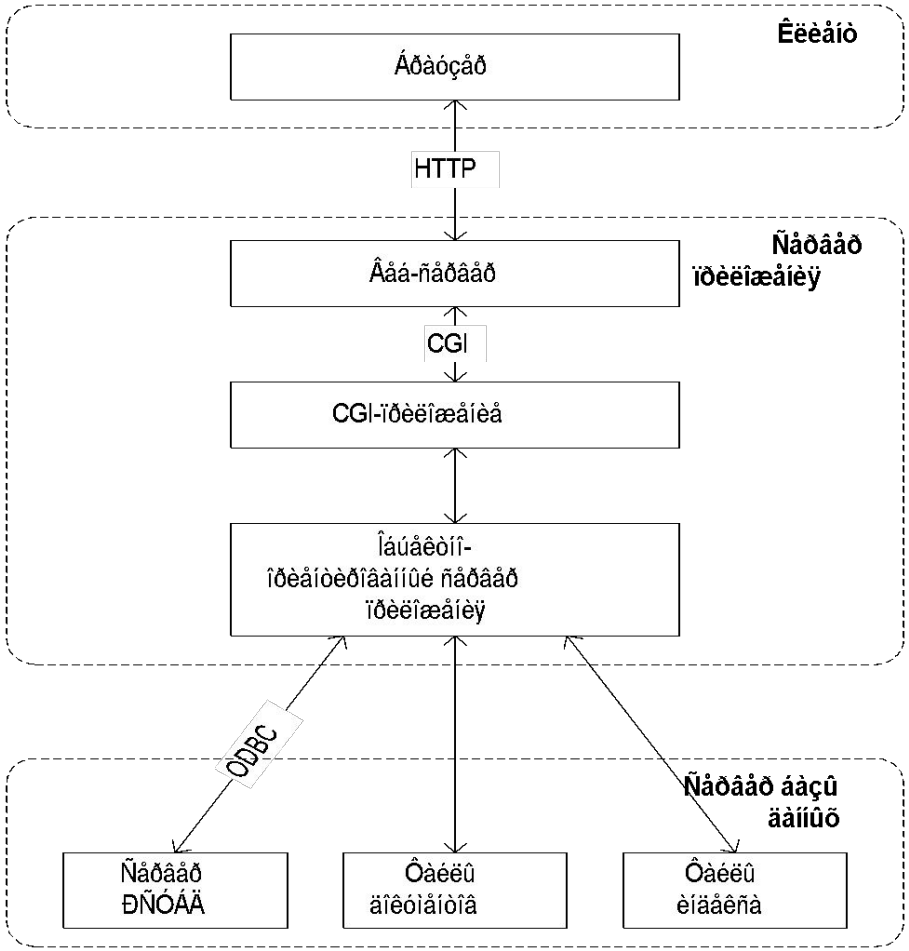
Основные ориентиры

- Объектно-ориентированная парадигма проектирования и программирования.
- Внешняя бизнес-логика (написанная на С++ и вынесенная за пределы конкретной СУБД в сервер приложения; многоплатформенность в плане независимости от СУБД);
- СУБД используется для хранения данных, обеспечения их транзакционной и ссылочной целостности;
- Сочетание прямого индексирования, необходимого для смыслового поиска, с построением инвертированных индексов, необходимых для быстродействия;
- Использование специализированных типов данных;
- Разработка и использование специализированных средств для презентации результатов поиска (в том числе – «человекоразмерная» визуализация многомерных данных).

Содержание

- Введение. От предоставления ресурсов – к многофункциональности и интеллектуальности.
- Многоцелевая информационная среда T-System, ее общая платформа T-Libra и специализированные блоки T-Archive, T-Media, T-Research, T-Education.
- Проблема интеллектуализации: ориентировочные этапы (от «объектов» к «агентам»).
- Проблема интеллектуализации: концептуальный, методологический и технологический аспекты.
- Методологические ориентиры для онтологии-I.
- Технологические ориентиры для онтологии-I.
- Заключение. О внедрении разработок и воспитании пользователя.
- Литература.

Технологические ориентиры для реализации онтологии-I: архитектура, организация ресурсов (1)



- Блок-схема трехзвенной Интернет-архитектуры T-System с внешней логикой (сервер приложений), взаимодействием через ODBC с сервером базы данных (для хранения данных) и специализированной системой индексации (для повышения эффективности поиска)

Технологические ориентиры для реализации онтологии-I: организация поиска и презентации его результатов (2)

Поиск и его результаты

- **Унифицированная для всей многоцелевой среды T-System настраиваемая каскадная поисковая система, включающая в себя**:
  - **механизм *нелинейных каскадных запросов***,
  - **формирующих *функциональные системы (структуры) пользовательского запроса* и соединяющих**
    - **результаты *полнотекстового поиска***,
    - **релевантные *тезаурусы и концепты* (со встроенной в них логикой развертывания знания),**
    - ***текстовые метаданные***,
    - **нетекстовые объекты различной модальности (двух- и трехмерная графика, звук, видео, анимация …),**
  - **и обеспечивающая «человеко» - или «машиноразмерную» *презентацию результатов запроса* (включая визуализацию данных).**

Заключение.
Внедрение разработок и воспитание пользователя

- **Внедрение многоцелевой ИС T-System в полном объеме и ее поэтапная интеллектуализация планируется в 2006-2010 гг. в Архангельском областном центре повышения квалификации специалистов культуры (АОЦПК),**
- **а также (фрагментарно) в рамках различных партнерских проектов с крупными библиотеками, архивами, музеями, академическими институтами, университетами.**

- **Внедрение такого рода многофункциональных информационных сред, ориентированных на *интеллектуальный информационный поиск в широкой социально-значимой сфере деятельности*, будет иметь важный положительный побочный социальный эффект –**

- **способствовать «воспитанию пользователя», - формированию класса пользователей, действующих не только по модели *«домохозяйки»* (что типично для поисковых машин Интернет), но и по более содержательным моделям, например, поведенческим моделям *«ученика»* или *«исследователя».***

Литература (1)

- **[1]** *Terry Smith, Alex Ushakov, Bill Heller*. Some Aspects of Developing and Using the Digital Learning Environment in Alexandria Digital Earth Prototype // Proc. of the 5th National Russian Research Conference “Digital Libraries: Advanced Methods and Technologies, Digital Collections” RCDL’2003, St.-Petersburg, Russia, 2003, p.18-25, http://rcdl2003.spbu.ru/proceedings/C1.pdf; *O.Agapova, R.Mayer, T.Smith, A.S.Ushakov, A.A.Ushakov, Stefan Decker*. Developing Digital Library Visual Services for Building a Lesson-Design Environment Prototype // Proc. of the 5th National Russian Research Conference RCDL’2003, St.-Petersburg, Russia, 2003, p.130-139, http://rcdl2003.spbu.ru/proceedings/C3.pdf.; *Ushakov A.S., Agapova O., Smith T., Gerber M., Ushakov A.A*. Exploring the Semantic Types of Relationships for Visual Query Development // Proc. of the 6th National Russian Research Conference RCDL’2004, Pushchino, Russia, Sept.19 – Oct. 1, 2004, p.141-149, http://www.impb.ru/~rcdl2004.
- **[2]** *Mary Marlino, Tamara Sumner*. A Model and Research Agenda for Educational Community-based Digital Libraries: The Digital Library for Earth System Education // Proc. of the 5th National Russian Research Conference “Digital Libraries: Advanced Methods and Technologies, Digital Collections” RCDL’2003, St.-Petersburg, Russia, 2003, p.26-34, http://rcdl2003.spbu.ru/ proceedings/C2.pdf.
- **[3]** *U.Schoepflin*. The Archimedes Project: Realizing the Vision of an Open Digital Research Library for the Study of Long-Term Developments in the History of Mechanics // Proc. of the 5th National Russian Research Conference “Digital Libraries: Advanced Methods and Technologies, Digital Collections” RCDL’2003, St.-Petersburg, Russia, 2003, p.124-129, http://rcdl2003.spbu.ru/proceedings/G2.pdf.

Литература (2)

- **[4]** *Марчук А.Г.* Электронные архивы, музеи и экспозиции // Труды 5-ой Всеросс. науч. конф. RCDL’2003, Санкт-Петербург, Россия, 2003. - Изд-во СпбГУ, 2003, с. 106-111, http://rcdl2003.spbu.ru/proceedings/E3.pdf.
- **[5]** *Ю.И.Шокин, В.А.Ламин, А.М.Федотов, В.Б.Барахнин, О.Л.Жижимов, Н.А. Мазов, Б.Н.Пищик, Н.Н.Покровский*. Распределенная информационная система «Виртуальный музей науки и техники СО РАН» // Труды 5-ой Всеросс. науч. конф. RCDL’2003, Санкт-Петербург, Россия, 2003. - Изд-во СпбГУ, 2003, с. 112-116, http://rcdl2003.spbu.ru/proceedings/E4.pdf.

Литература (3)

- **[6]** *С.Х.Ляпин, А.В.Куковякин*. Многофункциональная электронная библиотека T-Libra: WWS-архитектура, интегрированный каталог, настраиваемый мультирубрикатор, гибкий параметризируемый полнотекстовый поиск // Труды 5-ой Всеросс. науч. конф. RCDL'2003, Санкт-Петербург, Россия, 2003. - Изд-во СпбГУ, 2003, с. 292-299, http://rcdl2003.spbu.ru/proceedings/J4.pdf; *С.Х.Ляпин, А. В.Куковякин*. Виртуальная лаборатория для гуманитарных исследований на основе электронной библиотеки с гибким полнотекстовым поиском // Труды Всеросс. науч. конф. «Научный сервис в сети ИНТЕРНЕТ», г. Новороссийск (пос. Дюрсо), 20-25 сент. 2004 г. - М.: Изд. Московского гос. университета, 2004. - С. 45-47; *С.Х.Ляпин, А.В.Куковякин*. Концепт-ориентированный поиск в электронной полнотекстовой библиотеке с мультимодальным расширением // Труды 6-й Всеросс. науч. конф. RCDL'2004, Пущино, 29 сент. - 1 окт. 2004 г. - С. 127-134, http://www.impb.ru/~rcdl2004; *С.Х.Ляпин, А.В.Куковякин*. T-Media: от музейной библиотеки к информационной среде для интеграции музейных ресурсов и сервисов // Материалы 7-й ежегод. межд. конф. EVA 2004 Москва. - Москва, ГТГ, 29 нояб. - 3 дек. 2004 года. - М.: Изд. Центр ПИК Минкультуры России; Гос. Третьяков. Галерея, 2004, http://conf.cpic.ru/upload/eva2004/reports/doklad_57.doc // Материалы 7-й ежегод. межд. конф. EVA 2004 Москва. - Москва, ГТГ, 29 нояб. - 3 дек. 2004 года. - М.: Изд. Центр ПИК Минкультуры России; Гос. Третьяков. Галерея, 2004, http://conf.cpic.ru/upload/eva2004/reports/doklad_57.doc; http://www.adit.ru/rus/conference/adit2005/papers/paper.asp?nomer=4.; *С.Х.Ляпин, А.В.Куковякин*. Методология и технология создания многоцелевой информационной среды T-System на базе электронной библиотеки с гибким полнотекстовым поиском // Труды 7-ой Всеросс. науч. конф. RCDL'2005, Ярославль, Россия, 2005. - Изд-во ЯрГУ, 2005 (в печати).

Литература (4)

- **[7]** *В.Б.Тарасов*. От многоагентных систем к интеллектуальным организациям: философия, психология, информатика. – М.: Эдиториал УРСС, 2002. -352с.
- **[8]** *П.К.Анохин*. Избранные труды. Философские аспекты теории функциональной системы. М., 1978 / Философские аспекты теории функциональной системы, с. 27-48; Принципиальные вопросы общей теории функциональных систем, с. 49-106; Философский смысл проблемы естественного и искусственного интеллекта, с. 107-124; *П.К.Анохин*. Функциональная система как основа физиологической архитектуры поведенческого акта / П.К.Анохин. Избранные труды. Системные механизмы высшей нервной деятельности. – М., «Наука», 1979, с. 13-99.
- **[9]** *С.Х.Ляпин*. Функциональная организация субъекта и ценностные детерминации в научном познании // Ценностные детерминации в научном познании. Межвуз. сб. науч. статей. Отв. ред. д.ф.н., проф. Л.А.Микешина. – Вологда, 1984. – С.24-44.
- **[10]** Turchin,V.F. The Phenomenon of Science, Columbia University Press, 1977. (на русском: Турчин В.Ф. Феномен науки: Кибернетический подход к эволюции. Изд. 2-е - М.: ЭТС, 2000.http://www.ets.ru/turchin/Turchin,V.F. The Phenomenon of Science, Columbia University Press, 1977. (на русском: Турчин В.Ф. Феномен науки: Кибернетический подход к эволюции. Изд. 2-е - М.: ЭТС, 2000.http://www.ets.ru/turchin/); Heylighen, F., Joslyn, C., and Turchin, V. A short introduction to the Principia Cybernetica project, Journal of Ideas, Vol.2, No 1, pp.26-29, 1991. см.: Principia Cybernetica Project, http://pespmc1.vub.ac.beTurchin,V.F. The Phenomenon of Science, Columbia University Press, 1977. (на русском: Турчин В.Ф. Феномен науки: Кибернетический подход к эволюции. Изд. 2-е - М.: ЭТС, 2000.http://www.ets.ru/turchin/); Heylighen, F., Joslyn, C., and Turchin, V. A short introduction to the Principia Cybernetica project, Journal of Ideas, Vol.2, No 1,

Литература (5)

- **[11]** *В.Г.Редько*. Проблемы интеллектуального управления – общесистемные, эволюционные и нейросетевые аспекты (приглашение к круглому столу) // Институт прикладной математики им. М.В.Келдыша РАН, http://www.keldysh.ru/BioCyber/
- **[12]** *Л.М. Веккер*. Психические процессы. Т.2. Мышление и интеллект. – Л.: Изд-во Ленинградского университета, 1976. (Глава VI. Мышление как интегратор интеллекта, с.276-338, §2. Понятийная мысль как вид мышления и как форма работы интеллекта, с. 280-320); *Л.М.Веккер.* Психика и реальность: к единой теории психических процессов. - М.: Изд-во "Смысл", 1998; эл. версия http://www.psylib.ukrweb.net/books/vekkl01
- **[13]** *Заде Л.А*. Понятие лингвистической переменной и его применение к принятию приближенных решений. – М.: Мир, 1976. – 165 с.; *Заде Л.А*. Роль мягких вычислений и нечеткой логики в понимании, конструировании и развитии информационных / интеллектуальных систем (пер. с англ. И.З.Батыршина) // Новости искусственного интеллекта, №2-3, 2001, с. 7 - 11.

Литература (6)

- **[14]** *С.Х.Ляпин.* О концептах и концептологии (в поисках нового подхода к моделированию деятельности) // XIX World Congress of Philosophy. – Moscow 22-28 August 1993. – Book of abstracts. Сборник резюме. Vol. I. Секция 13 (Философия деятельности). – с.322; *С.Х.Ляпин.* Концептология: учение о концептах, методология культурогенных трансляций, технология эвристического развертывания смысла // Вестник СЗО РАО. – №3, 1998, СПб.-Архангельск: Поморский гос. университет им. М.В.Ломоносова, 1998. – с.28-41; *С.Х.Ляпин.* Концептологическая формула факта // Концепты. Научные труды Центроконцепта. – Отв. ред. С.Х.Ляпин. – Вып. 2(2) 1997. – Архангельск: Изд-во Поморского гос. университета, 1997. – с.5-71.
- **[15]** Ю.С.Степанов. Константы: Словарь русской культуры. Изд. 2-е. - М.: Академический проект, 2001.
- **[16]** *Н.Д.Арутюнова*. Предложение и его смысл: Логико-семантические проблемы. - М.: Наука, 1976; *Н.Д.Арутюнова.* Типы языковых значений: Оценка. Событие. Факт. – М.: Наука, 1988; *См. также*: Логический анализ языка. Сборник статей (1988 – 2002 гг). Отв. Ред. Н.Д.Арутюнова (и др.).
- **[17]** См. статьи «Текст», «Текстовой анализ», «Постмодернизм», «Постструктурализм», «Ризома» в кн.: Новейший философский словарь. Изд. 2-е, переработанное и дополненное. Сост. и гл. науч. ред. А.А.Грицанов. – Минск: Интерпрессервис. Книжный дом, 2001. – 1279 с.
- **[18]** *А.Н.Горбань, В.Л.Дунин-Барковский, А.Н.Кирдин, Е.М.Миркес, А.Ю. Новоходько, Д.А.Россиев, С.А.Терехов, М.Ю.Сенашова, В.Г.Царегородцев*. Нейроинформатика. – Новосибирск: Наука. Сибирское предприятие РАН, 1998. - 296с