

Холюшкин Ю.П.  
Витяев Е.Е.<sup>1</sup>,  
Костин В.С.

## Проблемы автоматизации информационных потоков в археологии

*Статья посвящена проблеме создания сетевых ресурсов, стимулирующих процессы становления и развития археологии как научной дисциплины.*

**Ключевые слова:** математизация науки, электронный документооборот, полевые отчеты, музейные технологии, библиотечные технологии, архивы, портал знаний.

### 1. Введение

С позиций системологии любая научная дисциплина в процессе развития проходит путь от простейших описаний предмета эмпирического исследования, через все большую детализацию, уточнение, специализацию собственного профессионального языка, до того состояния, когда становится возможным описывать изучаемые явления и процессы предельно конкретно и точно, используя языки все более абстрактных разделов математики.

Каждый шаг вперед на этом пути есть результат приобретения новой информации об изучаемом объекте: сами по себе математические методы не дают новой информации; они лишь позволяют преобразовать имеющуюся в постановке задачи информацию в такой вид, который более обозрим, удобен, полезен, пригоден для наших целей [Тарасенко, 2000: с.84].

Исходным понятием в этом ряду являются *технологии создания археологической информации*. Они охватывают все этапы и стадии формирования первичных данных, возникающих в процессе полевых, камеральных и кабинетных археологических исследований, включая работу с литературными источниками. Наиболее важными критериями целесообразности применения подобных технологий является полнота, достоверность и адекватность формируемых в исследовательском процессе данных. Все эти технологии в археологии разумно основывать на принципах и подходах, выделенных Л.С. Клейном [Клейн, 2001]. Среди них можно отметить следующие:

1. Принцип всесторонности, основанный на максимально возможном охвате информации.
2. Принцип комплексности, выраженный в преодолении тяги к уникальным вещам и поднятии интереса к массовому рядовому материалу.
3. Принцип универсальной подготовленности, согласно которому каждый археолог должен иметь представление о тех науках, специалистам которых он передает для анализа свой материал.
4. Стратиграфический принцип, под которым понимается возможность выявления относительной хронологии вскрываемых археологических материалов на основе их последовательности во времени.
5. Структурный принцип и принцип сопряженности. Первый из этих принципов подразумевает локальное взаиморасположение находок в археологическом комплексе, а второй – требует не разрознивать вещи из замкнутого комплекса ни в основной документации, ни в хранении, ни в публикации.
6. Полнота исследования – этот подход диктует требование тотального вскрытия памятника и сплошных раскопок местонахождений. Согласно этому подходу на стоянке, городище стало желательным раскопать всю их площадь. Из могильника взять не выборку погребений, а все – иначе не ясна структура могильника, не видно распределение разных видов погребений по его площади [Клейн, 2001: с. 76].
7. Принципы бережности и сохранения информации – на нем основана аккумуляция научных знаний и возможность обмена ими. При таком подходе главным требованием является фиксация информации записями, рисунками, чертежами и т.д. Из этого вытекает протокольный подход, который был впервые сформулирован Тэйлором в виде «хорошей» аксиомы для археологов: «важно не то, что ты нашел, а как ты нашел это», и излишне напоминать, что «как ты нашел это, можно установить только из твоих записей, но не из самих находок» [Клейн, 2001: с. 82].

<sup>1</sup> Работа выполнена при финансовой поддержке гранта РФФИ № 11-07-00560-а, интеграционными проектами СО РАН № 47, 115, 119, а также Советом по грантам Президента РФ и государственной поддержке ведущих научных школ, проект НШ-3606.2010.1

8. Принцип проверяемости. Это общеметодический принцип всякого исследования. Археолог не видел ни событий, ни общей картины исследуемого периода, а видел лишь источник, т.е. их отражение в вещи, отделенных от них бездной времени. И видел мелко, только один раз. Его наблюдение – раскопки неповторимо, т.к. памятник в результате раскопок исчез. Поэтому археологический отчет – это отражение отражения [Клейн, 2001].
9. Методичный подход предполагает изучение всех возможных разновидностей археологических артефактов и памятников без каких-либо произвольных пропусков.

Хотя традиционные формы и методы археологических исследований практически реализуют использование этих принципов и подходов, важное значение должно придаваться форме, в которой фиксируется археологическое знание. Здесь в первую очередь ставится задача использовать современные информационные технологии.

В любой научной организации рано или поздно появляется проблема «наведения порядка» в массиве документов, потому что в настоящее время ни одна из стадий кругооборота документов не обходится без бумажной волокиты. Для обеспечения целостности документов и упорядоченного ведения документации необходима система электронного документооборота, которая обеспечивает быстрый поиск необходимой информации и снижает вероятность случайного уничтожения документов, а также позволяет снизить расходы на канцелярские принадлежности и оргтехнику.

Ключевой характеристикой археологических учреждений, обосновывающей их принципиальное отличие от каких-либо промышленных или финансовых корпораций, является то, что обрабатываемая этими учреждениями информация формируется в результате тесного переплетения двух разнородных информационных потоков, которые условно можно охарактеризовать как «управленческий» и «научный».

## **2. Система электронного управленческого документооборота (СЭУД)**

СЭУД представляет собой совокупность технических и аппаратно-программных средств и технологий для создания и управления хранилищами документов.

Управленческие данные характеризуются высокой степенью структуризации. Они являются центральным звеном при организации административно-управленческой деятельности научных организаций. Эта система должна обеспечить среду, являющуюся основой для систем коллективной административной работы, учитывать специфику управляющих процессов организаций и интегрировать потоки информации в единую систему обработки описательной и числовой управленческой информации. Платформа системы электронного документооборота должна включать в себя необходимые инструментальные средства для обеспечения устойчивого функционирования СЭУД.

Автоматизация информационных потоков управленческой информации предполагает:

- а) внедрение доменной архитектуры локальной вычислительной сети археологического учреждения (интранет), средств централизованного администрирования и межсетевого взаимодействия;
- б) оптимизацию состава и расположения серверов с целью уменьшения нагрузки на оборудование и каналы связи;
- в) систему комплексных закупок необходимых технических и программных средств, организации серверного узла;
- г) решения по защите локальной вычислительной сети и данных от несанкционированного доступа из Интернет, определение политики информационной безопасности и их реализация;
- д) выбор системы хранения данных. В зависимости от объема информации и количества пользователей необходимо соответственно наличие на сервере одной из существующих СУБД: MySQL, MS ACCESS, MS SQL Server, ORACLE, IBM DB2 Database или Lotus Notes/Domino. Также возможно применение готовых специализированных программных комплексов типа 1С-управление. Выбор должен быть основан на представленном разработчиком технико-экономическом обосновании;
- е) внедрение систем коллективной работы;
- ё) интеграцию планово-финансовых, бухгалтерских, статистических и других отчетных данных;
- ж) к обязательным возможностям стоит относить: автоматическую регистрацию документов, ведение их регистрационных карточек, ведение номенклатуры дел, связывание документов, прикрепление файлов, работу со справочниками и словарями, разграничение прав доступа, наличие поиска по реквизитам и полному тексту;
- з) корпоративную электронную почту, и службу новостей;
- и) службы электронных конференций и интерактивного обмена должны обеспечить потребности общения сотрудников археологических учреждений и филиалов в режиме реального времени;
- й) предоставление средств электронной публикации информации о деятельности организации во внешний мир с использованием ресурсов и служб Web-сайта организации;

- к) обеспечение контролируемого и стандартизированного доступа к результатам административной и экспертно-аналитической деятельности институтов и их филиалов извне.
- л) IP-телефонию.

### 3. Система электронного научного документооборота (СЭНД)

В археологических организациях существенную часть корпоративных данных составляет научная информация в бумажной и/или электронной формах, поступающая наряду с «чисто» управленческими данными. Эта информация образует научный поток, в силу задач организации тесно интегрированный с управленческими данными. Научные данные являются, как правило, слабоструктурированными или неструктурированными.

Важнейшей процедурой в такого рода технологиях являются модели данных, регулирующих не только форму представления фиксируемых (вводимых) данных (тексты, рисунки, снимки, чертежи, таблицы и т.д.), но и те материальные (бумага, киноплёнка, аудио и видеокассеты, компьютерные средства) и логические (макеты данных) носители, на которых эти данные предусматривается размещать для их последующего использования. Для этой цели служат разнообразные системы управления базами данных и знаний (СУБД), в частности, ориентированных на гипертекстовую и мультимедийную форму представления информации.

Но для археологических учреждений существует ряд специфических проблем. Дело в том, что из представленных СУБД самой мощной и эффективной является ORACLE, которая требует наличия высокооплачиваемых опытных системных администраторов. Поиск вакансий в Internet показывает, что опытные программисты готовы работать как минимум за 25-30 тысяч рублей в месяц. Но всякая ли организация готова платить такие суммы?

Что касается более дешевого программного обеспечения для создания такого продукта, сейчас существует много бесплатно распространяемых средств разработки. Примером могут служить программные сборки Denwer и TopServer, в которые включены виртуальный сервер (Apache), MySQL (в TopServer еще и phpMyAdmin), поддержка языков php, java, html, css, perl, mysql и многое другое.

Это потребует меньших затрат на содержание системы, однако не освобождает организацию от приема на работу приходящих или штатных программистов<sup>1</sup>.

Для небогатых организаций археологического профиля наиболее доступным является внедрение СЭНД собственными силами, включая и сопровождение её с первых дней применения. В чем здесь преимущества?

*Во-первых, в СЭНД будут только те функции, которые действительно необходимы. Система не будет требовать больше ресурсов, чем это необходимо. Каждая археологическая организация сможет сама для себя определить необходимый набор функций.*

*Во-вторых, стоимость созданной своими силами СЭНД будет гораздо ниже готовой. Стоимость созданного продукта будет включать стоимость скачивания свободно распространяемых программ, зарплату разработчика и системного администратора с хорошими навыками работы на компьютере.*

*В-третьих, программный продукт будет настраиваем настолько, насколько это требуется. В тезаурусах будут лишь изначально используемые слова. Внешний вид форм и карточек будет изначально создан с учетом пожеланий будущих пользователей СЭНД.*

*В-четвертых, наличие тех или иных дополнительных возможностей будет зависеть только от времени, затраченного на создание или изменение СЭНД.*

*В-пятых, безопасность СЭНД будет всегда под контролем. Кроме того, к администратору будут при возникновении малейших неполадок или проблем обращаться сразу. Заодно станет ясно, надо ли изменять программный код, переобучать пользователей или закрывать слабое место в системе.*

*В-шестых, во время внедрения не потребуются дополнительных специалистов и капиталовложений.*

К этим шести большим достоинствам собственной разработки СЭНД можно сразу добавить недостаток: качество такого продукта будет зависеть исключительно от качества знаний, опыта и добросовестности принятого на работу специалиста, а также от правильного руководства процессом разработки. Но та экономия, которая получится в результате применения собственной СЭНД, может перекрыть все издержки на создание.

---

<sup>1</sup> В САТИ ИАЭТ СО РАН существовала наиболее дешевая в РФ система создания и обслуживания программных продуктов, основанная на использовании труда студентов ВКИ НГУ и ФИТ НГУ с приемом их на работу (с зарплатой от 1000 до 4000 руб.). При такой низкой зарплате естественной была высокая текучесть кадров, однако она компенсировалась непрерывностью процесса подготовки новых студентов, обеспечивающих преемственность в процессе разработки. Так, на вопрос одного из американских коллег на международной конференции: во сколько обошлось создание виртуального музея? – последовал ответ: 8080! Чего? Долларов? Евро? Что Вы – рублей, ответил автор доклада, чем поверг аудиторию в удивленно – шоковое состояние.

### 3. Проблемы внедрения СЭНД.

Сейчас настало время решить вопрос, могут ли создаваемые базы данных стать эффективным средством доступа к археологическим материалам. Ответ, по мнению Ж.-К. Гардена, однозначен: развитие «сетей данных» в разных науках рано или поздно затронет и археологию. Такие создаваемые общими усилиями «сети данных» должны включать в себя не только исходные материалы раскопок, но и результаты исследовательских работ [Гарден, 1983: с. 225]. Всё это позволит осуществлять информационный поиск не традиционным путем изучения отдельных трудов и периодики, хранящихся в библиотеках<sup>1</sup>, а прямым запросом к локальным или сетевым базам данных. Главное новшество, по мнению Ж.-К. Гардена, здесь не в применении электронной техники, а в возможности прямого доступа к полному тексту первоисточников, собранных либо в одном фонде, либо в распределенном виде, но не рассеянном по разным печатным изданиям, зачастую в урезанном тезисном виде. На это понадобится несколько десятилетий, чтобы отдельные исследователи и исследовательские учреждения приспособились к новой практике, которую большинство археологов пока еще не поддерживают [Гарден, 1983: с. 225-226]. При переходе к новым способам хранения археологической информации в сетях данных должны быть преодолены и другие отрицательные факторы: логические («зоопарк» форматов описания), экономические (затраты времени и другие препятствия доступа к информации) и психологические («жадность» в обобществлении результатов собственных исследований, нехарактерная для настоящих ученых) [Гарден, 1983: с. 225].

Для преодоления логического препятствия археологам страны необходимо начать совместные разработки по упорядочению понятий в предметной области археологии. Под предметной областью (ПО) понимается совокупность объектов, «рассматриваемых с точки зрения некоторого предмета исследования – совокупности существенных свойств (атрибутов) и отношений объектов исследования, описываемых в некоторой системе понятий предметной области. Предмет исследования может быть задан онтологией предметной области – специфицирующей в некотором формальном языке множество рассматриваемых объектов, связи между ними, систему понятий, и свойства объектов. Предмет исследования и онтология определяют “взгляд”, “точку зрения”, с которой рассматриваются (описываются в системе понятий) объекты предметной области, отношения и их свойства» [Витяев, 2010: с. 9-16].

Таким образом, онтология в археологии представляет собой систему, описывающую структуру определенной проблемной области, и состоящую из множества классов понятий, связанных отношениями, их определений и аксиом, задающих ограничения на интерпретацию этих понятий в рамках данной проблемной области. Онтология, как пример общего соглашения о семантике области, способствует установлению корректных связей между значениями элементов этой области, тем самым создавая условия для их совместного использования.

Суммируя приведенные выше определения онтологии, можно сказать, что онтология представляет собой точное описание (модель) некоторой части мира применительно к конкретной области интересов [Андреева, Сергеев, Холушкин, 2004, с. 39-44; Андреева и др., 2006: с. 832-840].

Таким образом, онтология – это четверка вида  $\langle C, D, R, A \rangle$ , где

C – множество понятий конкретной предметной или проблемной области;

D – множество определений понятий;

R – множество отношений (связей) между понятиями;

A – множество аксиом.

Попытки построения фрагментов онтологий в археологии уже предпринимались в виде унификации полевых работ [Медведев, Алаев, Алаева, 1978], классификаций артефактов [Медведев, 1981] и каталогизации музейных коллекций [Асеев, Поднозова, Шер, 1980]. Однако в то время они не получили широкого распространения, поскольку задача онтологизации науки еще не созрела. Сознвая эти трудности, Медведев писал: «создание хотя бы одного определенного массива терминологии разовым порядком с целью придания ему законодательного ранга невозможно» [Медведев, 1981, с. 18]. Идея интеграции научного документооборота в виде общих представлений была отражена в ряде публикаций сектора археологической теории и информации (САТИ) [см. Холушкин, Воронин, 2005]. По мере реализации были созданы информационные

<sup>1</sup> В настоящее время в России созданы распределенные корпоративные библиотечные системы, которые могут вести поиск по запросам заинтересованных пользователей.

ресурсы, описываемые ниже. При этом постепенно формировалось более четкое представление о необходимости интеграции отдельных ресурсов в единую систему документооборота.

Задачи автоматизации охватывают следующие области исследований:

#### 4. Автоматизация полевых отчетов

Большая часть разработок в этой области носит локальный характер. Среди разработок следует упомянуть следующие:

АИС «АРХЕОГРАФ», разработанный Стасом Васильевым (ИИМК РАН), представляет собой информационную систему для описания археологических памятников. Структура программы предусматривает регистрацию в едином государственном реестре памятников истории и культуры, учет и поиск информации, а также «организацию ее непосредственного взаимодействия с ГИС, с использованием топографических карт местности любого масштаба. Система состоит из нескольких информационных модулей: названия, категория, административная и географическая привязки, относительная и абсолютная хронология памятника, кадастр, библиография, архивные сведения и изобразительные материалы. Каждый модуль имеет свою структуру описания объекта» [<http://ardb.spb.ru/ru/about.htm>].

Рис. 1. Интерфейс БД «Памятники Кисловодской котловины» [Коробов, <http://www.archaeology.ru/ONLINE/Korobov/korobov.html>]

Помимо функций учета и хранения базовой информации о памятниках (археологические карты) и быстрого поиска необходимых данных, программа также включает функции научно-исследовательской системы, контроля и охраны археологических памятников (паспорта на археологические памятники и др.), вывода информации на карты в ГИС MapInfo и Google Map.

Однако в России подобные проекты все еще являются редкостью, а опубликованных материалов практически нет.

В качестве единичного примера можно привести работы Института археологии РАН в Кисловодской котловине [Коробов, <http://www.archaeology.ru/ONLINE/Korobov/korobov.html>].

База данных «Археологические памятники Кисловодской котловины» создавалась на основе реляционной системы управления базами данных Microsoft Access 97. Благодаря своим возможностям, этот программный продукт позволяет включать информацию количественного и качественного плана, заносить любой объем текстовой информации при помощи поля типа MEMO, а также присоединять любые объекты, поддерживающие протокол OLE. Имеется возможность создавать списки фиксированных значений признака, подставляемых в поле базы данных. Кроме того, в данном программном продукте можно осуществлять связь записей базы данных с графической информацией при помощи гиперссылок (рис. 1). Всего в БД используется более чем 1400 файлов фотографий, 230 планов и около 170 файлов рисунков.



При создании проблемно-ориентированной ГИС Д.С. Коробовым использовался программный пакет ГИС ArcView 3.0 [Афанасьев, Чернышев, 1997, с. 442-445]. В настоящее время отрабатывается методика изучения археологических памятников методами пространственного анализа.

#### а) Полевые отчеты в ИАЭТ СО РАН

Создание отчета (перевод полученных при раскопках материалов в электронную форму, максимально отражающую по полноте всю собранную документацию) – дело очень трудоемкое и занимает достаточно продолжительный период времени. Кроме того, часто не все материалы полевых исследований отражаются в отчетах. Прежде всего, это касается различных наблюдений или идей автора (ов) раскопок, которые отражены только в полевых дневниках. Еще 10 лет назад на интернет-форумах по археологической тематике активно обсуждался вопрос – нужен ли компьютер в поле, а уже при современном развитии информационных технологий и методики обработки материала выезд на полевые исследования без компьютеров кажется немыслимым.

Однако эффективная автоматизация становится возможной только после разработки информационной модели данных для упорядоченного хранения всех описаний и иллюстративных материалов, необходимых для генерации отчета. Пока не зафиксирован определенный формат хранения данных, задачи автоматизации рассыпаются, как дом, построенный на песке.

*Информационная модель отчета.* В тексте отчета должно быть представлено достаточно информации для того, чтобы специалист мог составить объективную картину проведенных работ и проверить справедливость выводов авторов раскопок. Для этого в нем должно найти отражение не только описание проведенных работ и извлеченного археологического материала, но и авторское видение обстоятельств исследования, составляющих ближайшее окружение эмпирического объекта, а также имевшихся в распоряжении исследователей материальных, финансовых и человеческих ресурсов [Костин, Постнов и др. 2010: с.60-65].

#### б) Разработка системы генерации полевых отчетов в ИАЭТ СО РАН

Первая попытка создания электронных полевых отчетов предпринималась И. Черниковым под руководством А.В. Постнова.

В 2011 году в САТИ стала разрабатываться система генерации полевых отчетов, в рамках дипломных работ студентов ВКИ НГУ (Рис.2).

В процессе выполнения проекта была предложена настраиваемая (гибкая) структура хранения описательной информации, включающей тексты, фотографии, карты и планы, локальные и GPS координаты элементов раскопок и находок. Такая архитектура программной системы позволяет отделить задачу создания содержательной структуры представления данных от задач разработки интерфейса и базы данных. В результате выделилась роль так называемого «архитектора» системы, который фактически самостоятельно разрабатывает систему, не обращая при этом к помощи программистов.

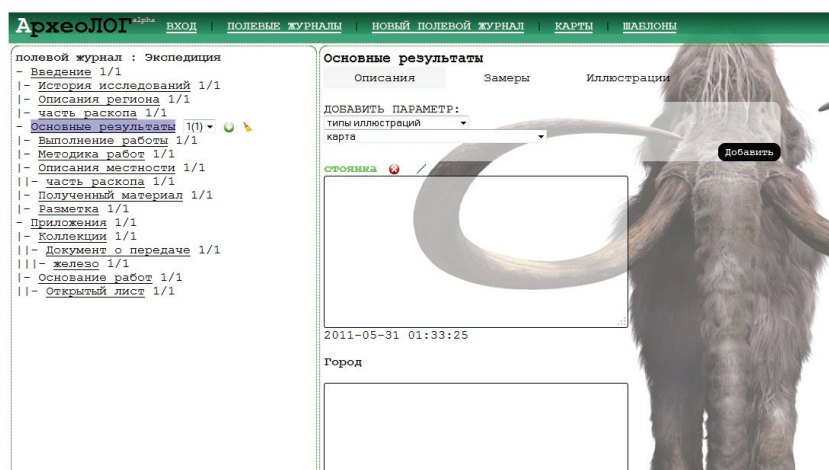


Рис.2. Страница формы полевого отчета

В результате разработки Web-приложения были решены следующие задачи:

- Разработка базы данных;
- Разработка пользовательского интерфейса;
- Разработка интерфейса для ввода географических координат;
- Разработка модуля генерации отчета.

Первая версия системы поддерживает генерацию полевых отчетов в виде HTML-документа.

Генерация отчёта предусмотрена в двух видах:

- По шаблону, разработанному «архитектором»;
- Полная (без шаблона) – для обзора всех описаний, относящихся к отчету.

В дальнейшем предполагается дополнить систему возможностями вывода в форматах документов MS Word, PDF и LaTeX.

### **Научный отчет**

Одним из направлений является разработка ресурсов, освобождающих научных сотрудников от рутинной работы. Создаваемая система в первую очередь предназначена для ученых секретарей научных подразделений Института археологии и этнографии СО РАН. Известно, что ежегодные научные отчеты отнимают много времени. Большая часть времени уходит на рутинную работу по сбору и перепроверке статистической информации (например, список сотрудников, их публикаций, экспедиционных поездок, участия в конференциях и т.д.).

Информация должна храниться из года в год и постоянно пополняться. При этом желательно организовать так, чтобы каждый сотрудник имел возможность добавить или изменить ту часть, которая касается лично его. Конечно, эта работа требует некоторой систематизации.

Можно попытаться хранить всю информацию в текстовых файлах (например, в формате Word или Excel, или же в каком-либо ином). Однако этот способ хранения данных может порождать характерные для подобных случаев проблемы. Например:

- проблему синхронизации при попытках внести изменения – один редактирует, остальные вынуждены ждать;
- необходимость соответствующего программного обеспечения для чтения и редактирования общих файлов;
- вольность использования форматирования разными людьми в силу разных представлений о конечном документе.

Стандартным решением таких проблем является отделение информации от формата ее представления, для чего идеально подходит инструментарий, включающий современные базы данных и способы доступа к ним по сети интернет. Современные инструменты манипулирования данными в сетевых базах данных позволяют их пользователям:

- выполнять одновременное и независимое редактирование отдельных записей;
- формировать итоговый отчет в произвольном формате.

Кроме того, в качестве доступа к информации, хранимой в базах данных, целесообразно использовать веб-интерфейс. Эта целесообразность основана на бесспорных преимуществах использования веб-интерфейса. Этот механизм доступа к данным:

- обеспечивает доступность с любого компьютера, подключенного в сеть;
- не требует установки дополнительного программного обеспечения.

Таким образом, наряду с решением перечисленных выше проблем ориентация на использование баз данных в подготовке и ведении научных отчетов дает возможность научным сотрудникам, их руководителям и ученым секретарями получать доступ к их ресурсам почти из любого места.

*Исходя из сложившихся традиций оформления и порядка ведения отчетов, в секторе археологической теории и информатики были вычленены основные составляющие статистической информации для отчета, которые мы будем в дальнейшем называть объектами:*

- гранты;
- конференции (рис. 3);
- экспедиции;
- публикации сотрудников сектора;
- международное сотрудничество;
- сотрудники сектора.

В силу некоторых ограничений реляционной базы данных в ней нельзя хранить объект произвольной структуры. Это означает, что каждое поле хранимого объекта должно быть простого типа (например, текст или число). Поэтому для хранения подобных объектов могут потребоваться дополнительные структуры, которые мы назовем подобъектами.

В подобъектах предполагается хранить те части объекта, которые не укладываются в простой тип. Например, список участников, соавторов, список работ и т.п. Подобъекты по своей структуре аналогичны объектам. Причем, если их структура не укладывается в простые типы, то они могут так же содержать подобъекты более низкого ранга.

В указанном выше случае все подобъекты оказались простого типа. В итоге были выявлены следующие объекты и подобъекты:

The screenshot shows a web browser window with the address <http://www.sati.archaeology.nsc.ru/intrabase/conference.html>. The page has a navigation bar with links: [гранты](#), [конференции](#), [экспедиции](#), [публикации сотрудников сектора](#), [международное сотрудничество](#), and [сотрудники сектора](#). Below the navigation bar is a section titled 'Конференции' with a sub-header 'добавление'. The form contains the following fields:

- \*участник: Холушкин Ю.П.
- \*название: 7-я ежегодная международная конференция EVA 2004
- \*статус: Международная (dropdown menu)
- \*место проведения: Москва
- \*начало: 29.11.2004 (calendar icon)
- \*окончание: 03.12.2004 (calendar icon)
- \*степень участия: докладчик (dropdown menu)
- \*тема доклада: Новые информационные технологии в археологии и этнографии
- содержание (ссылка): <http://www.tretyakov.ru>

At the bottom of the form are two buttons: 'Сохранить' and 'Отмена'. Below the form, there is a copyright notice: '© 2004, Сектор Археологической Теории и Информатики ИАЭТ СО РАН'.

Рис. 3. Пример заполнения формы по отчету об участии в научной конференции.

На основе выявленной структуры и ориентировочного содержания отчетов для каждого объекта и подобъекта были созданы таблицы в базе данных и сформированы скрипты для отображения списка уже существующих в базе записей, их редактирования и создания новых.

Полученная система позволяет вводить, редактировать и просматривать всю статистическую информацию. После добавления скрипта формирующего отчет на основе введенной информации и получаем готовую систему по хранению, редактированию и формированию отчета. [Холушкин, Воронин, Воробьев, 2004: с. 72-74; <http://www.sati.archaeology.nsc.ru/index.html?mi=otchet-y-z-a-period>] (Рис.3).

### Статистические исследования

В предыдущих разделах была рассмотрена схема построения сетевой среды – носителя информационных потоков, возникающих во время и вокруг археологических исследований. Они требуют применения сложных вычислительных методов (Рис.4).

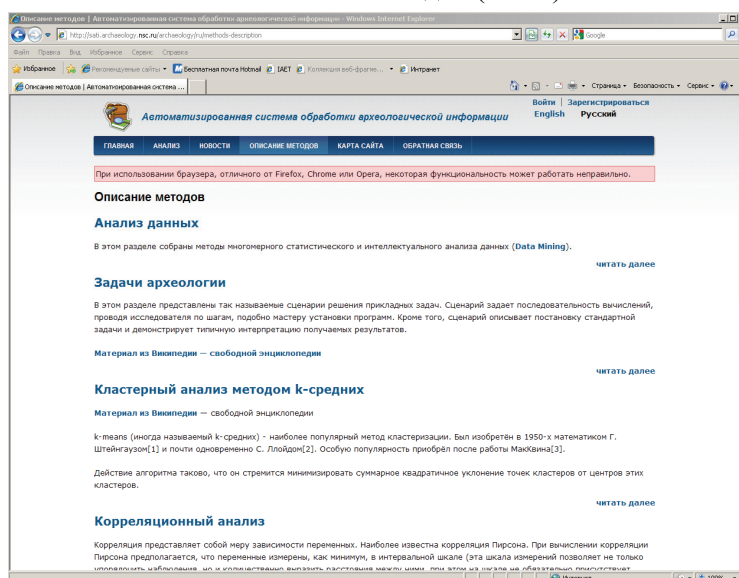


Рис.4. Страница описания методов в автоматизированной системе обработки археологической информации

Исследователям не всегда доступны такие вычислительные ресурсы в силу их дороговизны и сложности освоения. Отсюда возникает задача создания инструментария, который позволял бы:

1. объединить данные полевых исследований, проводимые в различных районах Сибири, в единую базу данных;
2. предоставлял простые в понимании средства для единообразного сбора, хранения, обработки и



- представления данных об археологических находках;
3. позволял через web интерфейс обрабатывать полученные данные специализированными, ориентированными на археологические данные методами анализа данных (см. след. статью)

**Электронный архив** — система структурированного хранения электронных документов, позиционируемая как основа документооборота, обеспечивающая надежность хранения, конфиденциальность и разграничение прав доступа к документам.

**Виртуальные музеи.** Электронные коллекции сосредоточены в Виртуальном музее истории и культуры народов Сибири и Дальнего Востока, создаваемом как система хранения, доступа и навигации по ресурсам музейных коллекций [Холушкин и др. 2008: с. 29-36].

Наполнение и доступ к информации электронного каталога музея осуществляется под управлением MySQL в операционной системе Linux. Обслуживание пользовательских запросов и генерация HTML-страниц производится с помощью программ на языке PHP, JavaScript, XML. Создаваемый Виртуальный музей истории и культуры народов Сибири и Дальнего Востока обеспечивает (Рис.5):

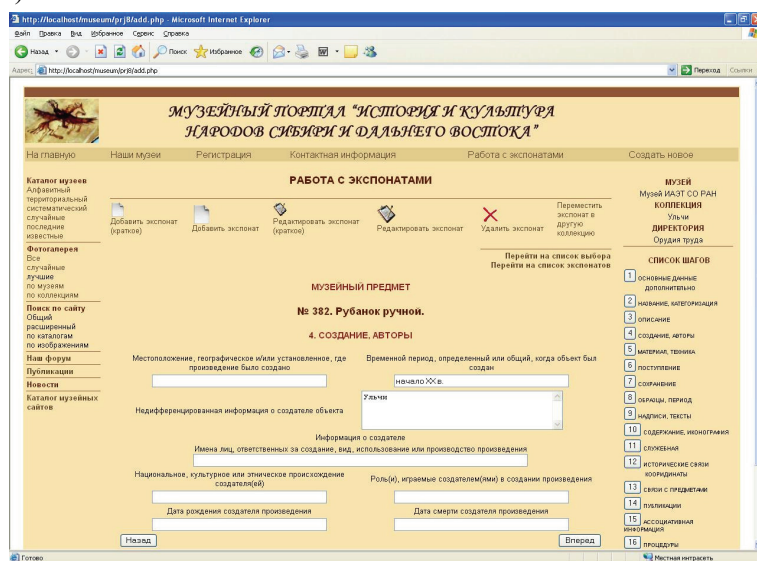


Рис.5. Полное редактирование музейных экспонатов.

- хранение и получение исчерпывающей информации о музейных экспонатах;
- соответствие стандарту международной организации по стандартизации ISO № (Z39.50) по хранению информации о культурном наследии – № (СИМІ);
- создание и хранение различного рода информации о музеях, виртуальных и реальных;
- сохранение изображения предметов, оцифрованных фотографий и набросков, так что каждый экспонат может хранить несколько изображений, и представления этих изображений в разных уровнях качества;
- просмотр их в удобных последовательностях;
- создание различных классификаций и типологических таблиц в виде директорий и удобную навигацию для них;
- быстрый поиск по самым "популярным" полям описания;
- расширенный поиск по любым полям формы описания;
- создание и просмотр публикаций и поддержку просмотра других публикаций, соответствующих музейной тематике и размещенных за пределами музея, и информацию о них;
- поддержку просмотра полезных ссылок, которые соответствуют музейной тематике, и информацию о них,
- поддержку просмотра последних новостей по музееведению и тп.;
- более полный доступ и расширенную информацию (например, по экспонатам) пользователю посредством назначения ему дополнительных привилегий;
- полный инструментарий по манипулированию различными категориями данных назначенному пользователю с предоставлением ему дополнительных привилегий;
- удобный и простой интерфейс;
- связь с VRML-версией музея и картографической системой.

Одним из приоритетных направлений развития музейного Web-сайта Института археологии и этнографии СО РАН и популяризации его музейных коллекций является активное использование технологий моделирования виртуальной реальности для поддержки полной иллюзии реального

свободного доступа в залы и хранилища музея для их посетителей и самостоятельного изучения находок и их описаний применительно к их вкусам и интересам. Компьютерные трёхмерные реконструкции, созданные по технологии виртуальной реальности и снабженные гипертекстовыми и мультимедийными структурами, могут помочь формированию существенно более целостного, наглядного и детального представления, дадут возможность погрузить музейные объекты в историко-культурный контекст, предоставят возможность самостоятельного обучения и исследования. Публикация таких моделей в сети Internet и/или на компьютере (т.н. музейном киоске) должна дать новую дополнительную возможность огромной аудитории значительно расширить свои представления о культурной и исторической ценности музейной экспозиции, а также о существующих в музее проектах реставрации.

**Электронная библиотека.** Сектором было в свое время осуществлено освоение и эксплуатация прикладных программ для наполнения пилотного варианта электронного каталога в базе данных научной библиотеки ИАЭТ СО РАН (археология и этнография) [Холушкин и др., 2003: с. 81-85]. При создании пилотного варианта в качестве программного обеспечения СУБД базы данных библиографических описаний была использована свободно распространяемое программное обеспечение СУБД WinIIsis. Для его освоения и последующего использования потребовалась локализация, так как кириллизированных версий этого программного продукта не существует. На основе разработанной участниками проекта кириллизированной версии программного обеспечения разработан единый формат данных электронного каталога и отработаны конвертеры (Рис. 6).

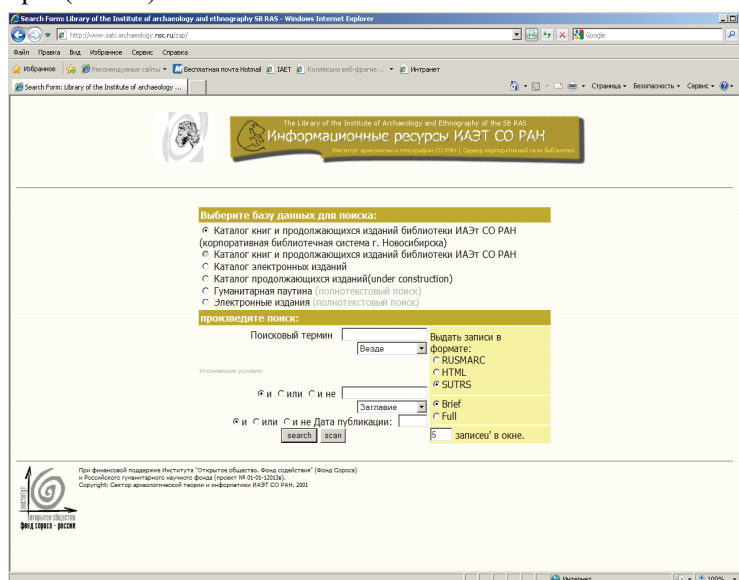


Рис.6. Главная страница электронной библиотеки

Последовательное применение формата RUSMARC при построении моделей данных электронного каталога позволяет представить все библиографические и иные описания в единой манере, что дает возможность строить соответствующие схемы конвертации данных из различных небольших по объему библиотечных информационных систем и сводить их в более крупные интегрированные. Доработан конвертер электронного каталога из формата базы данных под управлением СУБД WinIIsis в формат RUSMARC. Разработано и протестировано программное обеспечение Z-сервера (на Web-узле сектора археологической теории и информатики Института археологии и этнографии СО РАН) для последующего представления в Интернет пилотного варианта электронного каталога. Разработано и протестировано программное обеспечение Z-клиента (на Web-узле сектора археологической теории и информатики Института археологии и этнографии СО РАН) для доступа из Интернет к базам данных электронного каталога на Z-сервере. Разработан и протестирован Web-сайт с доступом к базам данных электронного каталога на Z-сервере сектора археологической теории и информатики Института археологии и этнографии СО РАН. Осуществлено опытное подключение электронного каталога научной библиотеки ИАЭТ СО РАН к корпоративной сети г.Новосибирска.

#### **Сибирский книжный археологический портал.**

Книжный портал реализован на системе управления содержимым сайта (CMS) с открытым исходным кодом WordPress, распространяемая под GNU GPL, написанной на PHP, в качестве базы

данных, используя MySQL. За основу дизайна была взята тема Twenty Ten, которая была адаптирована под потребности проекта "Сибирский книжный археологический портал". Проект рассчитан на заполнение группой людей, которые могут добавлять новые записи в соответствии со своими правами в существующую структуру (Рис.7).

Поддерживается разнообразный контент: текстовые файлы, doc-документы, pdf, большое количество форматов изображений, видео и аудио файлы.

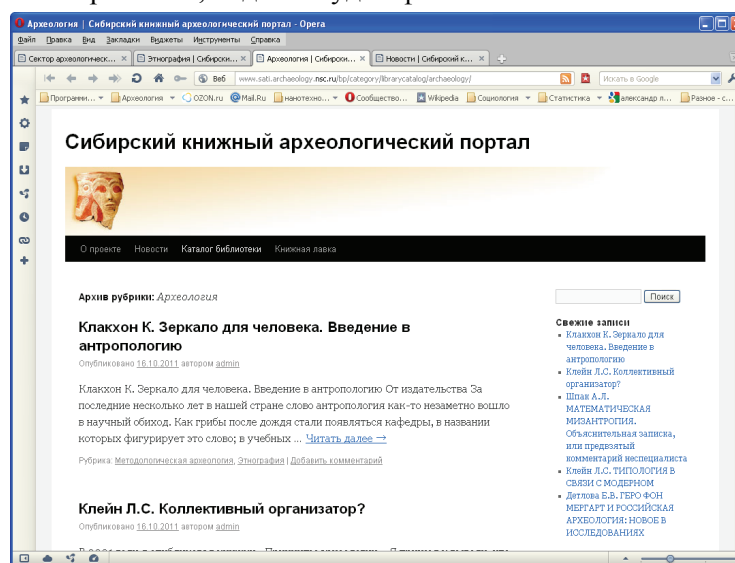


Рис. 7. Сибирский книжный археологический портал.

Права по изменению структуры могут быть делегированные нескольким пользователям. Так же имеется возможность модерации новых записей, в настоящий момент эта функция не используется. Имеется поддержка RSS, Atom. Возможен поиск статей по описаниям, как по всему portalу, так и по различным разделам.

### Электронные конференции

Каждый, кто подключился к сети Интернет, может посылать в ближайший сервер новостей свою информацию в виде статьи или произвольного двоичного файла, разместив ее в одной из рубрик. Через некоторое время (от одного до нескольких дней) ваша статья окажется на всех серверах новостей и будет доступна для всех пользователей сети Интернет. Если эта статья кого-либо заинтересует, возможно, вы получите на нее ответ. Этот ответ может быть либо опубликован в той же рубрике, либо прислан вам индивидуально по электронной почте. Вы, разумеется, можете тоже отвечать на любые статьи, опубликованные в любой рубрике [Фролов, Фролов, 1996].

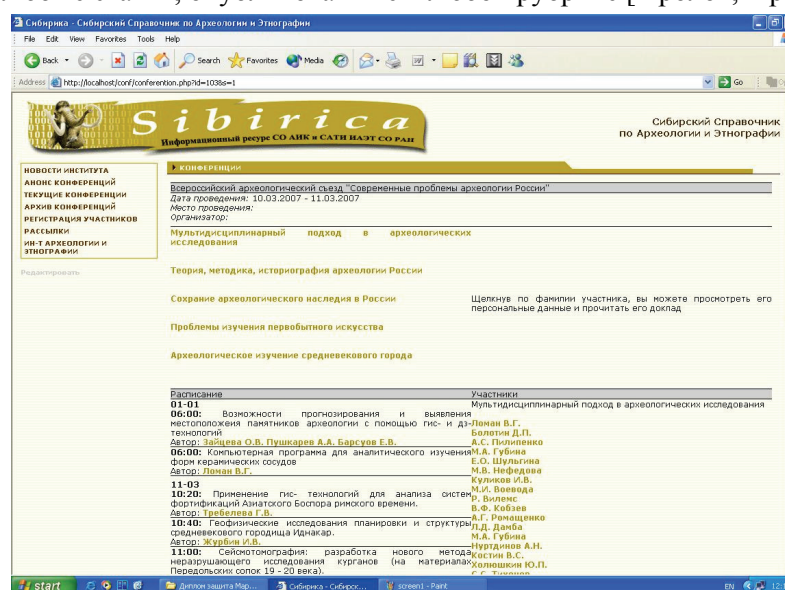


Рис.8. Страница просмотра информации о конференции:

Система электронных конференций чем то напоминает газеты. Каждая рубрика – это газета. Так же как и любая газета, рубрика имеет свое название, по которому можно догадаться о

содержимом. Вы можете выбрать любую "электронную газету" и прочитать любую статью из нее. Однако в отличие от обычных газет, статьи для которых пишут журналисты, система электронных конференций более демократична – каждый пользователь сети Интернет может опубликовать свою статью в любой "электронной газете", причем без предварительного согласия главного редактора.

Среди электронных конференций есть управляемые и неуправляемые. Первые находятся под контролем создавшего их человека, который определяет, публиковать вашу статью в конференции, или нет. Во вторых царит полная анархия. Вам могут также встретиться рубрики, закрытые для записи в них новых статей (Рис.8).

Посылая запрос специального вида, вы можете получить список всех рубрик, расположенных на сервере конференций, к которому вы подключены. Затем, опять же с помощью специального запроса, вы можете подписаться на одну или несколько заинтересовавших вас рубрик. После этого вам будут автоматически приходить все новые статьи из рубрик, на которые вы подписались.

Интерактивный способ предполагает наличие специальных программ, предназначенных для работы с электронными конференциями.

Средства работы с электронными конференциями встроены во все современные программные пакеты, предназначенные для работы с Интернет. Среди действующих электронных конференций можно назвать конференцию Северного конгресса.

### **Видеоконференции**

Технология, обеспечивающая передачу видеоизображения, звука и других типов данных на любые расстояния по цифровым каналам связи. Для вывода изображения могут быть использованы компьютерный или телевизионный монитор, плазменная или жидкокристаллическая панель, экран с мультимедиа-проектором и т.п.

Видеоконференция позволяет организовать многостороннее общение между учеными находящимися в разных городах и странах.

### **Порталы знаний**

Для решения задачи сведения выше приведенных археологических ресурсов в единое информационное пространство, обеспечения возможности открытого и удобного доступа к ним и поддержки их целостности нами предложена концепция специализированных Интернет-порталов знаний [Боровикова, Загорулько, 2002: с. 76-82]. На этой концепции основана разрабатываемая нами технология создания и сопровождения порталов знаний по гуманитарным наукам (Рис.9).

Информационную основу таких порталов знаний составляют онтологии, включающие как описание науки и научной деятельности в целом, так и описание конкретной научной дисциплины и соотнесенное с ним описание структуры и типологии соответствующих хранилищ данных и сетевых ресурсов [Боровикова, Загорулько, 2002: с. 76-82].

Благодаря предложенной структуризации системы знаний, когда явно выделяются предметно-независимые онтологии науки и научного знания, являющиеся общими для всех гуманитарных наук, портал знаний становится легко настраиваемым на выбранную предметную область. Так, при построении портала знаний для определенной гуманитарной дисциплины достаточно только построить ее онтологию и связать ее с предметно-независимыми онтологиями и соответствующими информационными ресурсами.

Для настройки портала на конкретного пользователя в состав информационной части портала включена модель пользователя, которая может постоянно уточняться и расширяться и тем самым всегда отражать актуальный "информационный портрет" пользователя.

В данной работе обсуждается основанный на предложенной концепции [Боровикова, Загорулько, 2002: 76-82] подход к разработке специализированного Интернет-портала, обеспечивающего содержательный доступ к систематизированным знаниям и информационным ресурсам по археологии.

С системной точки зрения, портал знаний представляет собой специализированную информационную систему, снабженную эргономичным пользовательским web-интерфейсом.

С точки зрения пользователя, портал является тематическим Интернет-ресурсом, обеспечивающим возможность поиска и просмотра информации в рамках заданной предметной области (гуманитарной дисциплины).

Как информационный ресурс портал:

- обеспечивает доступ к информации по различным аспектам и участникам научной деятельности, таким как: составляющие научной дисциплины (подразделы дисциплины, методы исследования,

используемые термины и понятия), персоналии исследователей, информация по группам, сообществам, организациям, включенным в процесс исследования;

- позволяет интегрировать близкие по тематике ресурсы, представленные в Интернет, и локальной сети;
- предоставляет средства поиска интересующей пользователя информации в рамках всего информационного пространства портала;
- обеспечивает информационную поддержку пользователей ресурса (например, анонсирование разного рода событий и мероприятий);
- поддерживает гибкий пользовательский интерфейс, позволяющий учитывать предпочтения пользователя по работе с ресурсом и предоставляемыми сервисами.

The screenshot shows a web browser window displaying the 'АРХЕОЛОГИЧЕСКИЙ ПОРТАЛ ЗНАНИЙ' (Archaeological Portal of Knowledge). The page features a navigation menu on the left and a main content area with a table of publications. The table has columns for 'Автор' (Author), 'Название публикации' (Publication Title), 'Дата публикации' (Publication Date), and 'Тип публикации' (Publication Type). The table lists various academic works, including books, articles, and conference proceedings, with authors' names and publication years.

Автор	Название публикации	Дата публикации	Тип публикации
Никольский, Сергей	50-я Региональная археологическо-этнографическая конференция «Полоса древности и современности» (Археологический отдел)	2000	статья
Тейлор, К.	A recently discovered province of southern Anatolia in the Hittite period	1999	статья
Тейлор, К.	A review of the archaeological evidence for food plants from the British Isles: an example of the use of the	1995	статья
Варенин, Василий	Archaeological Corridor Studies (ACCS)	2001	монография
Варенин, Василий	Analyse statistique des ensembles de Paléolithique moyen du Proche-Orient et du Caucase	2001	монография
Варенин, Василий	Archaeology and Post-Structuralism	2000	статья
Варенин, Василий	Centuries of Decline during the Holocene Classic Period at Puente de	2000	статья
Варенин, Василий	СНН, профиль 239.50 арх.объект, информация о культурном наследии	2008	статья
Варенин, Василий	Correlation des industries paléolithique moyen du Proche Orient et du Caucase	2004	монография
Варенин, Василий	СВЕРХ НОВЫЙ теоретический журнал	2004	журнал
Варенин, Василий	Современная археологическая наука (исследования)	1994	статья
Варенин, Василий	From the Editor	2002	статья
Варенин, Василий	Современная археологическая наука (исследования)	2002	статья
Варенин, Василий	ИЗДАТЕЛЬСТВО	2002	статья
Варенин, Василий	ИЗДАТЕЛЬСТВО	1995	статья
Варенин, Василий	Multiple comparisons in multiple response questionnaires	2000	статья
Варенин, Василий	Multiple comparisons in multiple response questionnaires	2000	статья
Варенин, Василий	On Modernity and Archaeological Discourse	2000	статья
Варенин, Василий	Palaeolithic Culture in Northern Asia	1994	монография
Варенин, Василий	Palaeolithic of North Asia and the Problems of Ancient Migrations	1990	монография
Варенин, Василий	Prehistory and Ethnology of Northeast Asia (course of lectures)	1992	монография

Рис.9. Страница публикаций портала археологических знаний

Описание предметной области основывается на системной классификации археологической науки, предложенной Ю.П. Холушкиным и Е.Д. Гражданниковым в [Холушкин, Гражданников, 2000: 58 с.] и развиваемой в настоящее время.

Системная классификация состоит из фрагментов определенной универсальной структуры. Стандартный классификационный фрагмент может быть представлен в виде семантической карты, которая служит геометрической моделью фрагмента. Расположение элементов фрагмента определяется позиционной и ранговой координатами, соответствующим критериям первичности – вторичности, антиэнтропийности – энтропийности и общности – частности понятий. Каждое понятие может давать начало фрагменту более низкого яруса, для которого оно служит фоновым понятием, т.е. данный фрагмент охватывает площадку данного понятия, располагаясь под ней.

Таким образом, геометрической моделью классификационной системы может служить трехмерное классификационное пространство, осями которого служат позиционная, ранговая и ярусная координаты.

В статье приведены некоторые предложения по созданию информационных потоков в археологической науке. Будут они востребованы или нет, покажет время.

## ЛИТЕРАТУРА

- Андреева О.А., Сергеев И.П., Холушкин Ю.П. Информационная система «Системная археология»// Информационные технологии в гуманитарных исследованиях. Выпуск 7. – Новосибирск: РИЦ НГУ, 2004, с. 39-44.
- Андреева О.А., Боровикова О.И., Булгаков С.В., Загорюлько Ю.А., Сидорова Е.А., Циркин Б.Г., Холушкин Ю.П. Археологический портал знаний: содержательный доступ к знаниям и информационным ресурсам по археологии // КИИ\_2006 Десятая национальная конференция по искусственному интеллекту с международным участием. Труды конференции. Т.3. – М.: Физматлит, 2006: с. 832-840.
- Асеев Ю.А., Поднозова И.П., Шер Я.А., Каталогизация музейных коллекций и информатика // Современный художественный музей. Проблемы деятельности и перспективы развития. – Л., 1980.
- Афанасьев Г.Е., Чернышев А.В. Применение ГИС-технологии в археологических исследованиях // Картография на рубеже тысячелетия. Доклады I Всероссийской науч. конф. по картографии. – М., 1997.
- Боровикова О.И., Загорюлько Ю. А.. Организация порталов знаний на основе онтологий. // Труды международного семинара Диалог'2002 "Компьютерная лингвистика и интеллектуальные технологии". Т.2. – Протвино, 2002: с.76-82.
- Витяев Е.Е. Извлечение информации из данных // Информационные технологии в гуманитарных исследованиях. Вып. 15. Новосибирск, 2010: с. 84-87.
- Гарден Ж.-К. Теоретическая археология. – М.: Прогресс, 1983: 296 с.
- Клейн Л.С. Принципы археологии. – СПб, 2001: 152 с.



- Костин В.С., Постнов А.В., Хасанов С.А., Диконская Е.К., Кисарова В.П. Подготовка и генерация отчета по археологическим полевым работам средствами информационных технологий: анализ проблем и постановка задач // Информационные технологии в гуманитарных исследованиях. Вып. 15. Новосибирск, 2010: с. 60-65.
- Медведев Г.И., К проблеме морфологического анализа каменного инвентаря палеолитических и мезолитических ансамблей Восточной Сибири // Описание и анализ археологических источников. – Иркутск, 1981.
- Медведев Г.И., Алаев С.Н., Алаева Т.В., Опыт применения полевой фиксации карточки на палеолитических местонахождениях // Нач.-техн. конф. Археология, этнография Восточной Сибири – Иркутск, 1978,
- Тарасенко В.П. Некоторые проблемы формализации гуманитарных знаний (на примере археологии) // Информационные технологии в гуманитарных исследованиях. Вып. 2. Новосибирск, 2000: с. 84-87.
- Фролов А.В., Фролов Г.В. Электронные конференции в Интернет // Газета "Капитал", рубрика "Три килобайта", 1996.
- Холюшкин Ю.П., Воронин В.Т., Воробьев В.В., Бердников Е.В., Федоров С.А., Жилицкая Г.Ю., Грищенко А.А., Лузин А.В. Электронный каталог научной библиотеки Института археологии и этнографии СО РАН (археология и этнография) // Информационные технологии в гуманитарных исследованиях. Выпуск 6. - Новосибирск: Редакционно-издательский Центр НГУ, 2003, с. 81-85.
- Холюшкин Ю.П., Воронин В.Т., Воробьев В.В. Информационная система по подготовке годовых научных отчетов // Информационные технологии в гуманитарных исследованиях. Выпуск 7. - Новосибирск: Редакционно-издательский Центр НГУ, 2004, с. 72-74 .
- Холюшкин Ю.П., Воронин В.Т. Сектор археологической теории и информатики: итоги десятилетия // Новосибирск: РИЦ НГУ, 2005, 26 с.
- Холюшкин Ю.П., Воронин В.Т., Ильиных М.Ю., Соловьева С.А. Разработка архитектуры системы музейного портала "История и культура Северной Азии и Дальнего Востока // Информационные технологии в гуманитарных исследованиях. Вып.12. - Новосибирск: ЗАО РИЦ "Прайс-Курьер", 2008, с. 29-36.

Беленький К.Г.,  
Витяев Е.Е.<sup>1</sup>,  
Костин В.С.,  
Холюшкин Ю.П.

## WEB-портал статистической обработки археологических данных

*Статья посвящена проблеме создания статистического пакета, предназначенного для обработки археологической информации. По замыслу авторов такой инструментарий предоставляет археологам простые для понимания единообразные средства представления, сбора, хранения и обработки данных археологических памятников. Описываемый пакет позволяет применять специализированные, ориентированные на археологические задачи методы анализа данных. Пакет также открывает возможность решать задачи археологии специально подготовленными последовательностями методов анализа данных (стратегиями решения задач). Добавление новых стратегий и методов в состав пакета не требует изменений программного кода сайта. Это позволит в дальнейшем интегрировать в состав пакета множество методов, свободно распространяемых по лицензии Open Source.*

**Ключевые слова:** Web-интерфейс, анализ данных, стратегии решения задач, статистический пакет, Drupal, PHP, MySQL, XML, нейронные сети.

### Введение

В свое время еще И. Кант сформулировал мысль, которую вслед за ним повторяли и интерпретировали многие философы, о том, что любая отрасль знания с тем большим основанием может называться наукой, чем чаще и успешнее она использует математику в собственных целях.

В предыдущей статье была рассмотрена схема построения сетевой среды – носителя информационных потоков, возникающих во время и вокруг археологических исследований. Вариант заполнения одного фрагмента этой схемы будет описан далее. Этот фрагмент включает хранение и обработку массивов, состоящих из описаний археологических находок (их информационных копий). Такая обработка требует применения сложных вычислительных методов.

Исследователям не всегда доступны такие вычислительные ресурсы в силу их дороговизны и сложности освоения. Другим недостатком существующих пакетов анализа данных является отсутствие в них инструментария для решения предметно-ориентированных задач.

Отсюда возникает задача создания инструментария, который позволял бы:

- объединить данные полевых исследований, проводимые в различных районах Сибири, в единую базу данных;
- предоставлял простые в понимании средства для единообразного сбора, хранения, обработки и представления данных об археологических находках;

<sup>1</sup> Работа выполнена при финансовой поддержке гранта РФФИ № 11-07-00560-а, интеграционными проектами СО РАН № 47, 115, 119, а также Советом по грантам Президента РФ и государственной поддержке ведущих научных школ, проект НШ-3606.2010.1