

Современный подход к построению информационно-аналитической системы состояния и развития научной сферы в вузах*

В статье изложены результаты анализа функционирующих на уровне вузов и федеральном уровне информационно-аналитических систем, описывающих научную сферу. Сделан вывод об отсутствии автоматической связи между информационно-аналитическими системами данных уровней и предложен подход к разработке архитектуры единой информационно-аналитической системы состояния и развития научной сферы в вузах.

Ключевые слова: научная сфера, информационно-аналитическая система, автоматизация.

A MODERN APPROACH TO THE CONSTRUCTION OF THE INFORMATION-ANALYTICAL SYSTEM OF STATUS AND DEVELOPMENT OF SCIENTIFIC SPHERE IN INSTITUTIONS OF HIGHER EDUCATION

The article presents the results of analysis of functioning at university and federal levels information-analytical systems describing the scientific sphere. It is concluded that there is no automatic link between the information-analytical systems of given levels, so an approach is suggested to architecting a single information-analytical system of state and development of scientific sphere in institutions of higher education.

Keywords: scientific sphere, information-analytical system, automation.

Введение

В настоящее время в Российской Федерации принята и реализуется концепция сосредоточения развития науки и инновационных идей в высших учебных заведениях. Воплощение в жизнь такого замысла требует не только собственно активизации научной и инновационной деятельности вузов, но и постоянного отслеживания такой деятельности, оценки ее результатов и прогноза ее эффективности.

Министерство образования и науки Российской Федерации в интересах реализации государственной научно-технической и инновационной политики осуществляет ежегодные мониторинги деятельности вузов:

- мониторинг оценки эффективности деятельности вузов;
- мониторинг результативности деятельности научных организаций, выполняющих научно-исследовательские, опытно-конструкторские и технологические работы;
- мониторинг перспективных научных, инновационных разработок, заделов, реализуемых научными, исследовательскими коллективами вузов и научных организаций.

Кроме того, все юридические лица (за исключением субъектов малого предпринимательства), результатом деятельности которых стали научные исследования и разработки, ежеквартально подают Статистический отчет формы 2-наука.

Наконец, раз в шесть лет все вузы проходят государственную аккредитацию.

Несомненным прогрессом в организации некоторых из этих масштабных мероприятий является возможность удаленного ввода информации в базы данных коллективного пользования в электронном виде, однако между собой эти базы никак не связаны и пользователям приходится фактически пять раз в год вводить одну и ту же информацию в разные базы данных.

С другой стороны информация, которую необходимо вносить во входные формы указанных выше мониторингов, зачастую не является первичной, а рассчитывается через другие показатели, учитываемые при организации текущей работы вуза. В

* Работа выполнена при финансовой поддержке Минобрнауки России в рамках базовой части государственного задания (код проекта 2976)



Наталья Владимировна Мамаева,

к.э.н., начальник управления

Тел.: (495) 411-66-33 (287)

Эл. почта: nvmamaeva@mesi.ru

Московский государственный
университет экономики, статистики
и информатики
www.mesi.ru

Natalia V. Mamaeva,

PhD (econ.), Head of Management

Тел.: (495) 411-66-33 (287)

E-mail: nvmamaeva@mesi.ru

Moscow State University of Economics,
Statistics and Informatics,
www.mesi.ru



Лев Борисович Милютин,

к.т.н., с.н.с., начальник отдела

Тел.: (495) 411-66-33 (307)

Эл. почта: lbmilyutin@mesi.ru

Московский государственный
университет экономики, статистики
и информатики
www.mesi.ru

Lev B. Milyutin,

PhD (tech.), Head of Department

Тел.: (495) 411-66-33 (307)

E-mail: lbmilyutin@mesi.ru

Moscow State University of Economics,
Statistics and Informatics,
www.mesi.ru

этой связи многие вузы создают свои базы данных и информационные системы, которые аккумулируют и обрабатывают именно первичную информацию. Одной из функций этих систем является формирование показателей, необходимых для ввода во входные формы мониторингов и других внешних отчетов.

Если рассматривать процесс сбора и анализа отчетной информации в аспекте его сквозной автоматизации, то в этой точке существует разрыв. Данные, полученные в автоматизированных системах учета и хранения информации (уровень вуза), приходится многократно вручную заносить в другие автоматизированные системы сбора, учета, хранения и анализа информации (уровень министерства).

Это классическая ситуация, решаемая современными информационно-аналитическими системами (ИАС) [1, 2].

1. Понятие и архитектура ИАС

Информационно-аналитическая система – это программно-реализованный комплекс, предназначенный для сбора, хранения и обработки разнородной информации, получаемой из различных источников, а также ее

анализа в целях обоснования необходимых управленческих решений.

В современных ИАС эффективное хранение информации обеспечивается наличием в составе ИАС различных источников данных. Задача обработки и объединения информации решается с помощью инструментов извлечения, преобразования и загрузки данных.

Анализ данных производится путем применения современных программных средств статистического анализа данных, моделирования и прогнозирования поведения управляемой системы, в нашем случае – научно-образовательной среды.

Архитектура информационно-аналитической системы в обобщенном виде может быть изображена с помощью рис. 1.

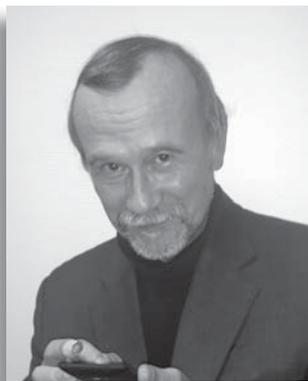
Данный рисунок демонстрирует этапы предварительной обработки, которые должны пройти первичные данные, прежде чем подвергнутся аналитической обработке.

Рассмотрим с позиции современных аналитических систем те программные средства сбора и анализа данных, на которых базируются указанные выше мониторинги.

На нижнем уровне (уровне вузов) мы видим наличие большого числа различных источников дан-



Рис. 1. Архитектура информационно-аналитической системы



Владимир Николаевич Николенько,
к.т.н., ведущий научный сотрудник
Тел.: (495) 411-66-33 (307)
Эл. почта: vnnikolenko@mesi.ru
Московский государственный
университет экономики, статистики
и информатики
www.mesi.ru

Vladimir N. Nikolenko,
PhD (tech.), Head Researcher
Tel.: (495) 411-66-33 (307)
E-mail: vnnikolenko@mesi.ru
Moscow State University of Economics,
Statistics and Informatics,
www.mesi.ru



Артём Игоревич Федосеев,
к.э.н., начальник отдела
Тел.: (495) 411-66-33 (307)
Эл. почта: aifedoseev@mesi.ru
Московский государственный
университет экономики, статистики
и информатики
www.mesi.ru

Artem I. Fedoseev,
PhD (econ.), Head of Department
Tel.: (495) 411-66-33 (307)
E-mail: aifedoseev@mesi.ru
Moscow State University of Economics,
Statistics and Informatics,
www.mesi.ru

ных. Это excel-таблицы, системы бухгалтерского и кадрового учета, базы данных и другие электронные средства хранения первичной информации, характеризующую деятельность научно-образовательных учреждений.

Разнообразие источников данных и необходимость их использования в каждом конкретном случае объясняется исторически сложившимися условиями, наличием собственной команды разработчиков в организациях, особенностями и спецификой научной деятельности организации. Если попытаться классифицировать источники данных по типам и назначению, то каждый из них можно условно отнести к одной из трех групп: транзакционные источники данных, хранилища данных, витрины данных [3].

Все источники данных первого уровня сбора и накопления первичной информации условно отнесем к понятию транзакционных источников данных.

Поскольку на этапе первоначального сбора данные, как правило, не согласованы друг с другом, то для дальнейшего анализа требуется их объединение и преобразование. Поэтому на следующем этапе должны решаться указанные задачи, а также устранение дублирования данных, после чего данные могут быть зафиксированы в аналитических базах данных. Аналитические базы данных и есть те основные источники, из которых в дальнейшем будет извлекаться информация для анализа с использованием соответствующих математических программно-реализованных методов.

При этом информационно-аналитическая система должна обеспечивать пользователям доступ к аналитической информации, защищенной от несанкционированного использования и открытой как через внутреннюю сеть организации, так и пользователям сети интранет и интернет. Архитектура современной информационно-аналитической системы насчитывает следующие уровни:

1. сбор и первичная обработка данных;
2. извлечение, преобразование и загрузка данных;
3. складирование данных;

4. представление данных в витринах данных;
5. анализ данных;
6. веб-портал [3].

2. Анализ существующих ИАС научной сферы Российской Федерации

Рассмотрим перечисленные уровни применительно к ИАС социально-экономического развития научной сферы Российской Федерации.

Сбор и первичная обработка данных

Первый уровень ИАС включает источники данных, ориентированные на сбор и обработку первичных данных, используемых в повседневной деятельности организации.

В последние годы большая часть университетов силами собственных разработчиков создает ИАС поддержки научной деятельности.

В качестве примеров следует упомянуть следующие системы, прошедшие достаточно длительные периоды практического использования:

- ИСТИНА – Интеллектуальная Система Тематического Исследования Научно-технической информации. Разработка МГУ [4];
- Информационно-аналитическая система сопровождения научно-исследовательской деятельности СПбГУ [5];
- Информационно-аналитическая система университета «Электронный университет» Казанского федерального университета [6];
- Информационно-аналитическая система «Наука». Самарского государственного технического университета [7];
- Информационно-аналитическая система Оренбургского государственного университета [8].

Необходимо отметить, что не все вузы используют в своей деятельности ИАС и учет данных по научной деятельности ведется в виде Excel- и Word-таблиц.

В любом случае речь идет о первичных данных, состав и форма представления которых отличается от тех данных, которые требуются для анализа на уровне министерства.

Итак, совокупность транзакционных источников данных образует нижнее звено архитектуры ИАС социально-экономического развития научной сферы. Целесообразно строить первый уровень указанной ИАС на основе уже имеющихся на вооружении систем сбора и первичной обработки данных, включающих транзакционные источники данных.

Извлечение, преобразование и загрузка данных

В современных ИАС извлечение, преобразование и загрузка данных осуществляются с помощью специально создаваемых ETL-инструментов (extraction, transformation, loading). Они содержат функции и протоколы извлечения данных из различных транзакционных источников нижнего уровня, их преобразования и консолидации, а также загрузки в целевые аналитические базы данных – хранилища данных и витрины данных. На этапе преобразования проводятся необходимые вычисления и устраняется избыточность данных.

Одним из центральных вопросов, решаемых на данном этапе, является определение функций и расчетных алгоритмов для показателей, которые используются министерством для оценки эффективности научной деятельности вуза, через первичные показатели, фиксируемые в транзакционных источниках данных.

Пусть в базе данных по учету кадров университета фиксируются следующие данные по сотрудникам:

F_i – фамилия сотрудника;

I_i – имя сотрудника;

O_i – отчество сотрудника;

D_i – должность сотрудника;

UZ_i – ученое звание сотрудника;

DR_i – дата рождения сотрудника.

Тогда показатель мониторинга оценки эффективности деятельности вузов «Численность работников вуза – руководители вуза – доктора наук – возраст 40–49 лет» будет рассчитываться по формуле

$n_d = \sum_{i=1}^N I_i$, если $UZ_i =$ «доктор» & $(40 < F(DR_i) <= 49)$, где $F(DR_i)$ – функция вычисления возраста по дате рождения.

Подобным образом через первичные данные можно определить все показатели, необходимые для мониторингов, проводимых министерством. Это будет также способствовать точности и однозначности терминологии.

После этого можно сформулировать транспортные протоколы для автоматической передачи данных на следующий уровень иерархии ИАС.

Складирование данных

К третьему уровню архитектуры ИАС относятся «хранилища данных». Они включают в себя источники данных, предназначенные для хранения и анализа информации. Такие источники объединяют информацию из нескольких транзакционных систем и позволяют анализировать ее в комплексе с применением современных программных инструментов анализа данных. В нашем случае здесь должны храниться данные всех подведомственных организаций, охваченных системой мониторингов Министерства образования и науки Российской Федерации.

Автор идеи складирования данных Б. Инмон определяет хранилище данных как предметно-ориентированную, интегрированную, некорректируемую, зависимую от времени коллекцию данных, предназначенную для поддержки принятия управленческих решений.

Характерными особенностями хранилищ данных являются: относительно редкая корректируемость большинства данных, обновляемость данных на периодической основе, единый подход к поименованию и хранению данных вне зависимости от их организации в исходных источниках.

Хранилище данных, являясь одним из главных звеньев архитектуры ИАС, выступает в качестве основного источника данных для всестороннего анализа всей имеющейся на данном уровне архитектуры ИАС информации.

Представление данных в витринах данных

Четвертый уровень архитектуры ИАС – это «витрины данных». Они предназначены для проведения

целевого анализа. Витрины данных строятся, как правило, на основе информации из хранилища данных.

С точки зрения пользователя, отличие витрин данных от хранилища данных заключается в том, что хранилище данных соответствует уровню всего министерства, а каждая витрина обычно обслуживает уровень отдельных департаментов и отделов, отличаясь достаточно узкой целевой специализацией.

Отличие витрин данных от транзакционных баз данных заключается в том, что первые служат для удовлетворения потребностей конечных пользователей, не являющихся профессиональными программистами: аналитиков, менеджеров разных уровней, решающих различные задачи бизнеса. Транзакционные же базы данных используются в основном операторами, отвечающими за ввод и обработку первичной информации, а не за ее анализ, нацеленный на поддержку принятия решений.

Применение витрин данных, многомерных и реляционных, в сочетании с современными инструментами делового анализа данных позволяет превратить просто данные в полезную информацию, на основе которой можно принимать эффективные решения.

Анализ данных

Следующий уровень архитектуры ИАС аккумулирует современные программные средства, позволяющие министерству проводить всесторонний анализ информации. Они помогают успешно ориентироваться в больших объемах данных, анализировать информацию, делать на основе анализа объективные выводы и принимать обоснованные решения, строить прогнозы, сводя риски принятия неверных решений к допустимому минимуму.

Инструменты интеллектуального анализа данных используются аналитиками для доступа к информации, ее визуализации, многомерного анализа и формирования как predeterminedных по форме и составу, так и произвольных отчетов. В качестве входной информации для делового анализа выступают заранее обработанные данные из

хранилища или представленные в витринах данных.

Следует учитывать, что ИАС социально-экономического развития научной сферы является уникальной системой, призванной обеспечить решение достаточно специфической задачи эффективного управления развитием научной сферы. В силу своей специфики инструментарий анализа данных, наряду с универсальными методами: статистический анализ данных, OLAP (*online analytical processing*, аналитическая обработка в реальном времени) и другими, должен содержать достаточно мощный блок математического моделирования и прогнозирования социально-экономического развития научной сферы [9].

Веб-портал

Логика развития архитектуры ИАС с неумолимостью привела к широкому использованию интернет-технологий. Сайты и веб-порталы постепенно приобретают все более весомую роль в архитектуре ИАС.

Возможность использования обычного веб-браузера для доступа к информации позволяет экономить на затратах, связанных с закупкой и поддержкой настольных аналитических приложений для большого числа клиентских мест. Реализация веб-портала позволяет снабжать аналитической информацией пользователей и аналитиков в любой точке мира, подключенных к порталу через интернет.

Мониторинг научных организаций проводится с использованием профессионально проработанного веб-портала. Пользователи, пройдя регистрацию, получают возможность удобного ввода исходной информации в систему, а по завершении ввода информации сразу же видят результаты ее аналитической обработки. На рис. 2 показан пример одного из результатов аналитической обработки информации, полученный через веб-портал мониторинга.

Аналогичные возможности получают пользователи веб-портала мониторинга оценки эффективности деятельности вузов. Отличие в том, что результаты аналитической обработки информации отображаются на

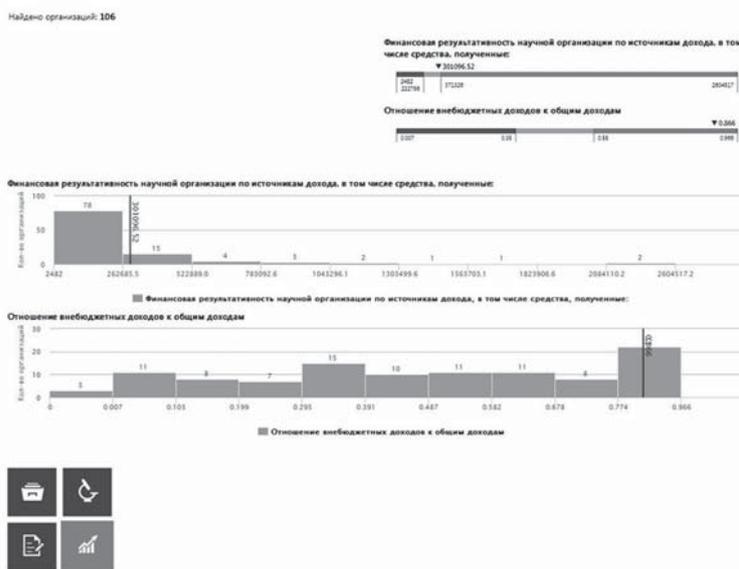


Рис. 2. Результат аналитической обработки информации из хранилища данных «Мониторинг научных организаций», полученный через веб-портал мониторинга

Наименование показателя	Значение показателя	Пороговое значение	Прогнозируемое пороговое значение
Образовательная деятельность	55,95	60	61,43
Научно-исследовательская деятельность	375	51,28	73,61
Международная деятельность	0	1	1,34
Финансово-экономическая деятельность	9404,55	1327,57	1893,65
Инфраструктура	221,45	13,92	15,09
Приведенный контингент студентов	33,8	220	264,00
Дополнительный показатель	2,29	2,78	3,34



Рис. 3. Результат аналитической обработки информации из хранилища данных «Мониторинг оценки эффективности деятельности вузов», полученный через веб-портал мониторинга

портале только после даты официального объявления результатов.

На рис. 3 показан пример одного из результатов аналитической обработки информации, полученный через веб-портал мониторинга.

Однако несмотря на то, что результаты анализа данных стали доступны в удаленном режиме для любого зарегистрированного в системе пользователя, цели такого анализа далеко не прозрачны. Если вся система мониторингов строится только для того, чтобы выявить «неуспешные» вузы, то она не стоит тех средств, которые затрачены на ее разработку.

На наш взгляд, информационно-аналитическая система социально-экономического развития научной сферы должна быть, прежде всего, инструментом для научного обоснования управленческих решений. Эти решения должны вырабатываться на основе детального анализа данных о развитии научной сферы, определения тенденций этого развития, математического моделирования процессов развития и анализа последствий принятия таких решений [10].

Ввиду чрезвычайной сложности задач анализа, моделирования и прогноза информация, накапливаемая в хранилище данных, должна

быть доступной для зарегистрированных пользователей и необходимый инструмент целесобразно разрабатывать по принципу открытых систем силами научного сообщества вузов.

Заключение

Проведенный авторами статьи анализ показывает, что рассмотренные министерские программные системы несут в себе основные черты, присущие современным информационно-аналитическим системам, однако у них есть и определенные недостатки:

1. Отсутствие возможности автоматического извлечения и преобразования первичной информации из транзакционных вузовских источников данных, таких как системы бухгалтерского учета, системы учета кадров, базы данных, Excel-таблицы. Фактически это означает отсутствие нижнего уровня системы и интерактивный ввод данных непосредственно в хранилище данных.

2. Наличие нескольких информационно-аналитических систем, обрабатывающих фактически одну и ту же первичную информацию, что заставляет вузовские коллективы многократно проделывать

работу по заполнению различных отчетов на основе одной и той же первичной информации.

3. В части анализа данных не вполне отчетливо определены цели анализа (кроме сокращения числа вузов). Непонятно, на каких моделях базируется прогноз развития вузовской науки, как оцениваются социальные и экономические последствия управленческих решений.

Авторы предлагают следующий подход к построению информационно-аналитической системы социально-экономического развития научной сферы.

Описать состав информации, хранимой в едином хранилище данных. Каждую переменную, фиксируемую в хранилище данных, определить через набор первичных данных и разработать точный алгоритм ее вычисления.

Организовать сбор и фиксацию данных в едином хранилище информации на основе принципа сквозной автоматизации. Для этого необходимо реализовать транспортные протоколы и программные средства автоматической выгрузки данных, рассчитанных на основе первичной информации с помощью разработанных алгоритмов, например, в транспортные Excel-табли-

цы, и автоматическую загрузку их в хранилище данных.

На уровне хранилища данных организовать несколько витрин данных для нужд различных групп пользователей информации, в том числе различных мониторингов, статистических ведомств, департаментов, а также вузов, зарегистрированных в системе.

Инструментарий для анализа данных дополнить блоками математического моделирования и прогноза, которые следует создавать как открытую систему, привлекая к этому процессу научную общность вузов.

При оценке научных показателей вузов учитывать специфику и профиль научно-образовательной деятельности, например нужно по-разному оценивать объем привлеченных финансовых средств для вузов гуманитарной, естественнонаучной, технической сферы, особенно сферы искусства.

По мнению авторов, такой подход позволит, во-первых, избежать неоднозначности толкования понятий, во-вторых, исключить неоправданные затраты на неоднократный ручной ввод информации в систему, в-третьих, повысить оперативность и качество анализа и прогноза.

Литература

1. Белов В.С. Информационно-Аналитические Системы. Основы проектирования и применения: учебное пособие, руководство, практикум / Московский государственный университет экономики, статистики и информатики. – М., 2005.
2. Говорков А.С. Автоматизация организационно-управленческих аспектов научной деятельности вуза // Университетское управление. – 2009. – № 6. – С. 13–18.
3. Волков И., Галахов И. Архитектура современной информационно-аналитической системы // Директор ИС. – 2002. – № 3.
4. ИСТИНА – Интеллектуальная Система Тематического Исследования Научно-технической информации [Электронный ресурс] / МГУ. – Режим доступа: http://www.chem.msu.ru/rus/events/khokhlov_261012.pdf
5. Информационно-аналитическая система сопровождения научно-исследовательской деятельности СПбГУ [Электронный ресурс]. – Режим доступа: <https://ias.csr.spbu.ru/>
6. Информационно-аналитическая система университета «Электронный университет» Казанского федерального университета [Электронный ресурс]. – Режим доступа: <http://kpfu.ru/dis/informacionno-analiticheskaya-sistema-kfu>
7. Информационно-аналитическая система «Наука». Самарский государственный технический университет [Электронный ресурс]. – Режим доступа: <http://www.systemworld.ru/node/56>
8. Информационно-аналитическая система Оренбургского Государственного Университета [Электронный ресурс]. – Режим доступа: <http://ecsocman.hse.ru/univman/msg/18456300.html>
9. Мамаева Н.В. и др. К вопросу о математическом моделировании социально-экономических процессов научной сферы // Актуальные проблемы развития науки и образования: сб. научных тр. по материалам Международной научно-практической конференции 5 мая 2014 г.
10. Тимофеева Н.Л. и др. Этапы создания модели научной сферы, как инструмента прогнозирования развития науки и технологий // Материалы конференции. Международная научно-практическая конференция «Ценности и интересы современного общества». Экономика и управление. Часть 2 // Московский государственный университет экономики, статистики и информатики. – М., 2014.