

**Ф.О. Федин,  
Т.В. Морозова,  
Е.Н. Павличева**

## **Модель информационно-аналитической системы обработки данных малых инновационных предприятий при высших учебных заведениях**

В статье излагается модель информационно-аналитической системы обработки данных малых инновационных предприятий, функционирующих в сфере образования и науки. Представление модели выполняется с использованием графического языка описания бизнес-процессов в нотации IDEF0 (функциональная модель). Для моделирования используется специализированное CASE-средство фирмы Computer Associates — AllFusion Process Modeler. Дается краткая характеристика основных составных частей модели, входных и выходных данных, управляющих и исполнительных механизмов.

*Ключевые слова:* информационно-аналитическая система; малое инновационное предприятие; ETL-процесс; хранилище данных; подсистема Business Intelligence; правила аналитической обработки данных.

**В**недрение малых инновационных предприятий (МИП) в высших учебных заведениях является пока еще развивающимся направлением преобразований в сфере образования и науки, получившим реальную юридическую основу с выходом Федерального закона № 217-ФЗ от 02.08.2009.

Как показывают исследования, создание МИП на базе высшего учебного заведения предоставляет этому образовательному учреждению такие преимущества, как дополнительное финансирование, возможность продвижения инновационных проектов, стимулирование научной деятельности. Студенты, обучающиеся по направлениям подготовки вуза, получают возможность (в том числе и в рамках прохождения производственных практик) формирования реальных практических умений и навыков решения практических задач будущей профессиональной деятельности. Происходит общее повышение рейтинга высшего учебного заведения [2; 5].

Опыт функционирования МИП при вузах и НИИ показывает, что в процессе своей деятельности управленческий персонал МИП решает такие многочисленные аналитические задачи, как нахождение неочевидных правил поведения клиентов и определение степени их лояльности, выявление неизвестных закономерностей спроса и предложения, прогнозирование тенденций развития производства и потребления, и другие. Наиболее полный перечень аналитических задач представлен в работе [2].

Анализ выводов исследований, описанных в работе [3], позволяет заключить, что все эти задачи могут с успехом решаться аналитическими инструментами, которые либо должны входить в состав автоматизированных информационных систем вузов, либо быть самостоятельными информационно-аналитическими системами малых инновационных предприятий (ИАС МИП). В качестве подхода к построению одного из таких инструментов авторами предложена модель ИАС МИП, разработанная с использованием CASE-средства «AllFusion Process Modeler». На рисунке 1 показана контекстная диаграмма функционирования этой модели.



Рис. 1. Контекстная диаграмма A-0 уровня работы ИАС МИП

В качестве входных данных в модели используются данные СУБД (например, «Access» или «SQL Server»), учетных систем (например, «Галактика» или «1С:Предприятие»), локальных документов (например, MS Excel или Word), внешних источников (например, веб-служб), различные архивы данных. При этом на вход ИАС МИП должны поступать не только непосредственно данные функционирования МИП, а и метаданные — структурированные, кодированные данные, описывающие характеристики объектов-носителей информации, способствующие идентификации, обнаружению, оценке и управлению этими объектами.

Управляющими параметрами (правилами, стратегиями, стандартами и др.) для модели являются: ограничения и правила формирования моделей аналитической обработки данных; правила формирования или модификации процессов функционирования моделей; существующие стандарты выполнения аналитических запросов и технологии Data mining; требования к системе хранения данных (хранилищу данных Data warehouse); регламенты ETL-процесса — извлечения данных из различных источников, их дальнейшей трансформации и загрузки в хранилище данных; правила выполнения аналитической обработки данных. Механизмами выполнения работ модели служат приложения и администраторы процессов ETL и Business Intelligence (BI), а также администраторы систем управления базами данных (СУБД).

Выходные параметры модели ИАС МИП — это либо данные, полученные в результате извлечения и визуализации, либо результаты применения специально построенных и обученных аналитических моделей. При этом извлечение и визуализация используется в тех случаях, когда для получения новых знаний об исследуемом объекте нет необходимости строить сложные модели, достаточно «посмотреть» на данные в нужном виде (подключить другой визуализатор, например, представить данные не в табличном виде, а в виде гистограммы), чтобы сделать определенные выводы о характере зависимостей и тенденций, получить ответ на интересующий вопрос.

На рисунке 2 изображена декомпозиция контекстной диаграммы первого уровня, на которой отражены основные структурные элементы ИАС МИП — ETL-процесс [1], многомерное хранилище данных, подсистема аналитической обработки данных.

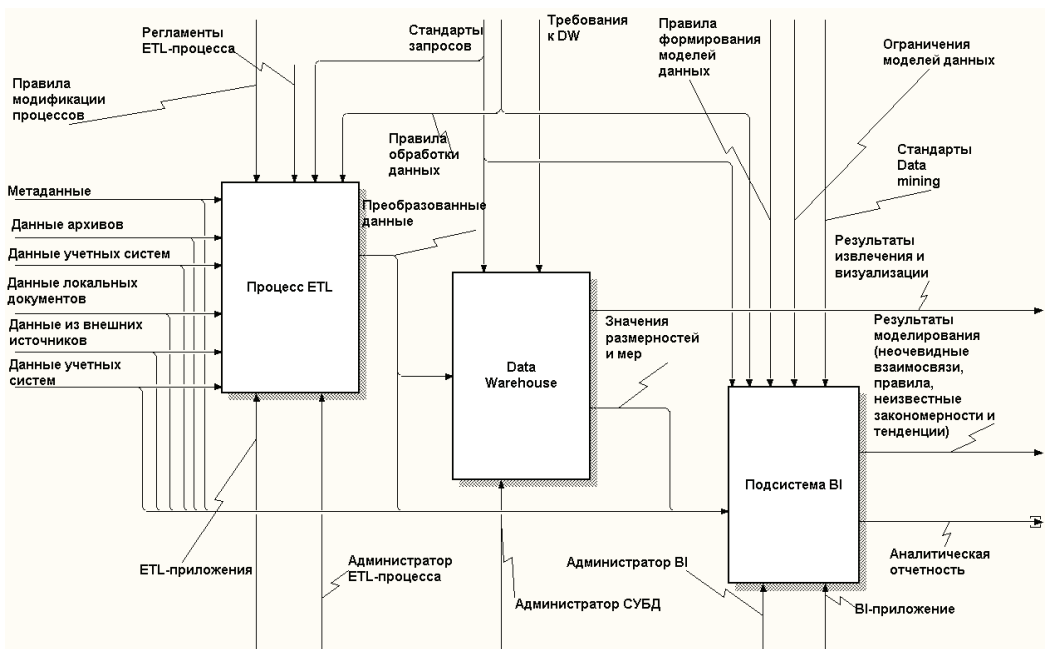


Рис. 2. Декомпозиция контекстной диаграммы первого уровня в нотации IDEF0

В рамках ETL-процесса (рис. 3) обеспечивается извлечение, преобразование и загрузка данных в систему хранения. Извлекаемые данные помещаются в промежуточные таблицы, после чего решаются вопросы их преобразования (трансформации). Необходимость преобразования данных обусловлена тем, что данные, поступающие на вход модели ИАС МИП, собираются из источников, имеющих разные типы. Эти данные создавались с применением разных программных приложений, при этом использовались различные методологии и стандарты. Структурная организация таких данных может существенно отличаться от структуры целевых таблиц хранилища данных ИАС МИП, что не позволит выполнить их загрузку в это хранилище. Следовательно, эти данные необходимо преобразовать с использованием специальных процедур. Если преобразование невозможно, то такие данные должны загружаться в специальные таблицы исключений.

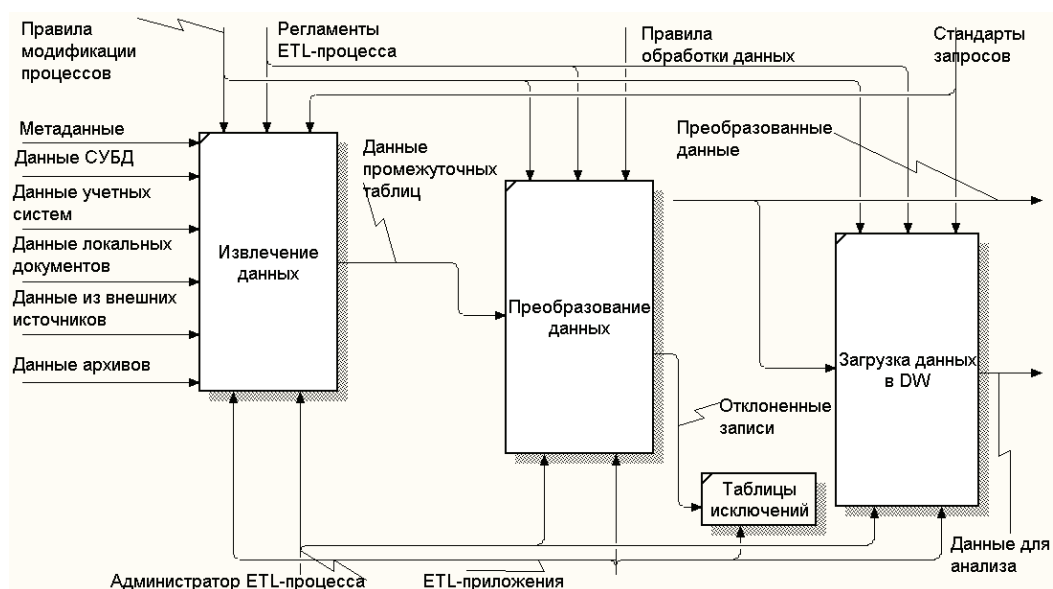


Рис. 3. Декомпозиция ETL-процесса

Преобразование данных включает: преобразование структуры данных, агрегирование данных, перевод значений, создание новых данных и очистку данных. Возможные варианты и методы преобразования данных подробно описаны в работе [4].

Подсистема Business Intelligence (BI) (рис. 4) предназначена для решения задач аналитической обработки данных МИП на основе построения и применения различных аналитических моделей. В ней используются методы поиска ассоциативных правил и последовательных шаблонов, регрессионного и кластерного анализа, классификации.

В качестве управляющих параметров при построении таких моделей выступают правила формирования или модификации процессов функционирования

моделей, существующие стандарты выполнения аналитических запросов и технологии Data mining, установленные ограничения.

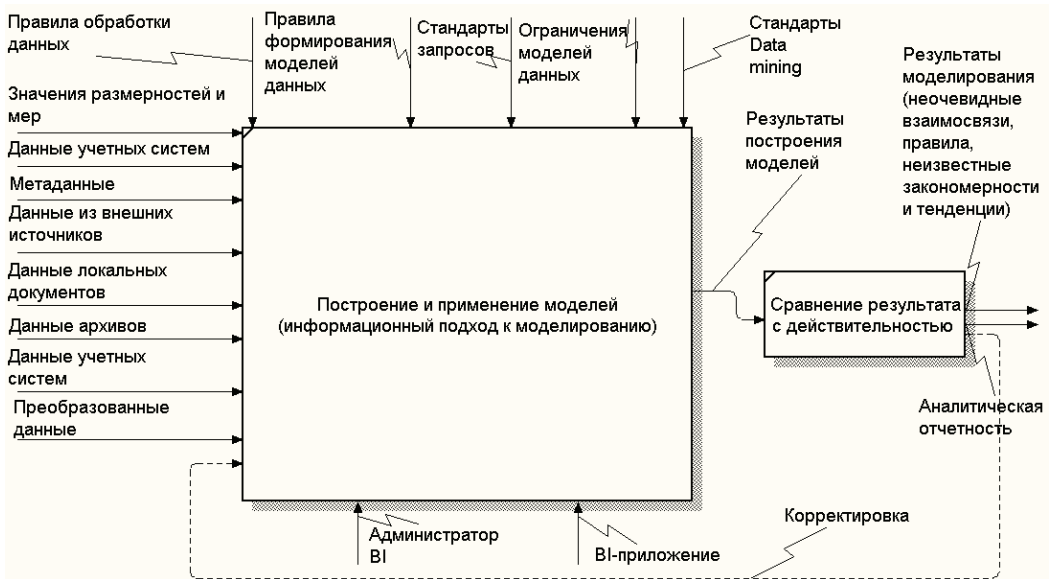


Рис. 4. Декомпозиция подсистемы Business Intelligence

Управляющими и исполнительными механизмами процессов разработки и использования аналитических моделей служат администраторы процесса Business Intelligence и BI-приложения, например, аналитическая платформа Dedactor studio, разработанная компанией BaseGroup Labs [1].

Процесс построения моделей включает шаги формулирования целей моделирования, подготовки и сбора исходных данных, непосредственно разработки (подбора) модели и ее обучения (с учителем и без учителя [4]) и проверки ее качества. Здесь важно отметить, что при построении аналитических моделей должен применяться исключительно информационный подход к моделированию, при котором вначале исследуется характер представленных данных и общая задача анализа, а затем происходит адаптация разработанной модели [1]. При таком подходе характеристики модели в полном объеме определяются имеющимися входными данными, следовательно, она не может быть построена без наличия исходных данных. Для корректировки параметров модели предусматривается обратная связь (рис. 4). Сам процесс настройки модели, как правило, носит итеративный характер.

Таким образом, в статье представлена разработка функциональной модели информационно-аналитической системы обработки данных МИП при высших учебных заведениях, разработанная с использованием графического языка описания бизнес-процессов в нотации IDEF0. На основе этой разработки могут быть созданы логические и физические модели модулей систем аналитической обработки данных МИП.

*Литература*

1. Паклин Н.Б., Орешков В.И. Бизнес-аналитика: от данных к знаниям. СПб.: Питер, 2009. 704 с.
2. Межевов А.Д., Федин Ф.О. Исследование основных задач аналитической обработки данных сферы образования // Вестник Университета (Государственный университет управления). 2011. № 24. С. 232–236.
3. Федин Ф.О. Возможности аналитических модулей в информационных системах образовательных учреждений // Вестник Московского городского педагогического университета. Серия «Информатика и информатизация образования». 2013. № 1 (25). С. 119–126.
4. Федин Ф.О., Федин Ф.Ф. Анализ данных: учеб. пособие. Ч. 1: Подготовка данных к анализу. М.: МГПУ, 2012. 204 с.
5. Федин Ф.Ф., Павличева Е.Н., Федин Ф.О. Критерии оценки инновационной деятельности вуза ИТ-профиля // Вестник Российского университета дружбы народов. Серия «Информатизация образования». 2013. № 2. С. 93–99.

*Literatura*

1. Paklin N.B., Oreshkov V.I. Biznes-analitika: ot danny'x k znaniyam. SPb.: Piter, 2009. 704 s.
2. Mezhevov A.D., Fedin F.O. Issledovanie osnovny'x zadach analiticheskoy obrabotki danny'x sfery obrazovaniya // Vestnik Universiteta (Gosudarstvenny'j universitet upravleniya). 2011. № 24. S. 232–236.
3. Fedin F.O. Vozmozhnosti analiticheskix modulej v informacionny'x sistemax obrazovatel'ny'x uchrezhdenij // Vestnik Moskovskogo gorodskogo pedagogicheskogo universiteta. Seriya «Informatika i informatizaciya obrazovaniya». 2013. № 1 (25). S. 119–126.
4. Fedin F.O., Fedin F.F. Analiz danny'x: ucheb. posobie. Ch. 1: Podgotovka danny'x k analizu. M.: MGPU, 2012. 204 s.
5. Fedin F.F., Pavlicheva E.N., Fedin F.O. Kriterii ocenki innovacionnoj deyatel'nosti vuza IT-profilya // Vestnik Rossijskogo universiteta druzhby narodov. Seriya «Informatizaciya obrazovaniya». 2013. № 2. S. 93–99.

*F.O. Fedin, T.V. Morozova, E.N. Pavlicheva*

**Model of Information and Analytical System of Data Processing  
of Small Innovative Enterprises at Universities**

The paper sets out the model of information and analytical system of data processing for small innovative enterprises operating in the field of education and science. Presentation of the model is performed using a graphical language for describing business processes in the notation IDEF0 (functional model). For modeling a specialized CASE-tool of the firm Computer Associates — AllFusion Process Modeler is used. A brief description of the main components of the model, the input and output data, controlling mechanisms and actuators.

Keywords: information and analytical system; small innovative enterprise; ETL-process; data warehouse; subsystem Business Intelligence; rules of analytical processing.