

**К ВОПРОСУ РАЗРАБОТКИ ХРАНИЛИЩА ДАННЫХ ДЛЯ АНАЛИЗА  
ТУРИСТИЧЕСКОЙ ОТРАСЛИ РФ**

**THE ISSUE OF THE DEVELOPMENT DATA WAREHOUSE FOR  
ANALYSIS THE TOURISM INDUSTRY OF THE RUSSIAN FEDERATION**

**Н.А. ГЕРАСИМОВ,**

доцент кафедры бизнес-аналитики НИУ-ВШЭ, кандидат технических наук,  
г. Москва, [gera01@yandex.ru](mailto:gera01@yandex.ru)

**N.A. GERASIMOV,**

Higher School of Economics, Department of Business Informatics School, Associate  
Professor, Candidate of Technical Sciences, Moscow, [gera01@yandex.ru](mailto:gera01@yandex.ru)

**Б.Е. МАГИН,**

научный сотрудник кафедры рекреационной географии и туризма  
географического факультета МГУ имени М.В. Ломоносова, г. Москва,  
[maginboris@gmail.com](mailto:maginboris@gmail.com)

**B.E. MAGIN,**

Lomonosov Moscow State University, Faculty of Geography, Department of  
Recreational Geography and Tourism, Research Assistant, Moscow,  
[maginboris@gmail.com](mailto:maginboris@gmail.com)

**Аннотация.**

В работе разработан вариант подхода к определению нового класса компьютерных моделей, ориентированных на сложную аналитику. Такие модели образуют класс Complete Analysis System (CAS). На примере модели анализа туристического бизнеса России показано, как модели класса CAS расширяют функциональность базовой структуры аналитической системы, в основе которой лежит модель «многомерного хранилища». Разработка хранилища макро показателей туристического бизнеса в концепции многомерного «Куба» позволяет значительно повысить эффективность инструментальной части аналитической модели и применить гибкий OLAP-режим для многовариантных исследований туристической отрасли РФ. Предложенный подход возможно применить и для других сложных экономических приложений, реализованных в концепции Complete Analysis.

**Ключевые слова.**

Бизнес-аналитика, полный анализ, комплексный анализ, хранилище данных, многомерное хранилище, модель, туризм, туристический бизнес, OLAP-системы.

### **Abstract.**

A variant of the approach to the definition of a new class of computer models, focused on complex analytics. These models form a class Complete Analysis System (CAS). On the example of Russian tourist business model analysis, it is shown as a model CAS class extend the functionality of the basic structure of the analytical system, which is based on the model of “multi-dimensional storage”. Development of storage macro indicators of tourist business in the concept of a multi-dimensional “Cube” can significantly improve the efficiency of the instrumental analytical model and apply flexible OLAP-mode for the multivariate study of the tourism industry of the Russian Federation. The proposed approach may be applied to other complex economic applications implemented in the Complete Analysis of the concept.

### **Keywords.**

Business analyst, a full analysis, complex analysis, data warehouse, multidimensional storage model, tourism, tourist industry, the OLAP-system.

Начиная с 2000-х годов на рынке международных туристических услуг наблюдается устойчивый годовой прирост (около 4 %) [1]. Расширяются как традиционные виды туризма (пляжный, курортный, медицинский, религиозный, которые увеличились примерно на 30 %), так и его новые виды: сельскохозяйственный, экологический, экстремальный, познавательный, винный и др. Для таких государств, как Греция, Турция, Испания, Египет и Израиль туризм является важным доходным бизнесом, составляющим в бюджете страны значительную долю.

Очевидно, наблюдаются и некоторые изменения в структуре направлений туризма. Так, еще несколько десятилетий назад основную долю занимал «купально-пляжный» туризм или туризм с целью отдыха по классификации UNWTO (Всемирная туристическая организация ООН), а в настоящее время идет существенное увеличение (примерно до 30 %) доли туристов, посещающих родственников; туристов, ищущих качественное медицинское обслуживание или осуществляющих путешествия с религиозными целями [1]. Например, количество туристов, путешествующих с целью получения медицинского обслуживания, по данным Deloitte Touche Tohmatsu Limited [2], вырос за 5 лет примерно в 3 раза и составил около 11 миллионов человек в мире, при этом средние траты в поездке составляют от 4 до 15 тысяч долларов

США. Одним из факторов, влияющим на резкий рост туристов, путешествующих с целью получения медицинских услуг в других странах, является активное старение населения во многих развитых странах, что приводит, с одной стороны, к росту расходов у населения на медицинское обслуживание, а с другой стороны, ведет к сокращению отчислений в различные медицинские фонды работающего населения, так как непрерывно изменяется структура занятых.

Следует отметить, что Россия также идет в тренде роста как выездного, так и внутреннего туризма. Раскрываются новые возможности для организации туристического отдыха в Крыму, на Камчатке, Дальнем востоке, на Алтае. Создаются новые туристические маршруты в центральных районах России.

Однако серьезным фактором, сдерживающим темпы развития туристической отрасли в России, является отсутствие достаточной информированности по туристическим ресурсам нашей страны и слабая информационная поддержка туристического бизнеса со стороны государства.

Так, например, при анализе данных по медицинскому туризму нельзя не отметить некоторые факторы, оказывающие сдерживающее действие на рост отрасли в целом. В первую очередь это плохая информированность о медицинских услугах в стране проживания и за рубежом, сложность выбора между различными медицинскими сервисами и отсутствие единых критериев и правил предоставления услуг, а также практическое отсутствие на рынке медицинского туризма глобальных бизнес-игроков.

Дальнейшее развитие туристического бизнеса в России (и в частности, медицинского туризма), по мнению авторов, невозможно без применения современных информационно-коммуникационных технологий, которые позволят создать на базе мощных хранилищ данных современные инструменты для информирования, анализа, построения прогнозов и управления отраслью в целом. Такие технологии требуют разработки и применения единых стандартов и критериев к участникам туристического бизнеса, а также должны предлагать

типовые (или стандартные) высокопродуктивные технические решения организации единого информационного обеспечения туристической отрасли.

Одним из возможных проектов в решении данной проблемы может стать единая система сопровождения *информационного пространства* туризма России, которая должна включать информационно-справочную и информационно-аналитическую подсистемы.

Основной задачей информационно-справочной подсистемы является предоставление достоверной и легкодоступной информации о различных направлениях, центрах и видах туризма. Пользователь такой информационной системы должен получить достоверную и качественную информацию о ценовых, ресурсных и рекреационных возможностях того или иного туристического центра (или маршрута), иметь возможность сравнить туристический потенциал и обменяться мнением с потенциальными пользователями туристического ресурса. По существу, такая информационно-справочная система играет роль глобальной «социальной сети» потребителей и производителей различных продуктов туристического бизнеса.

Попытки создания подобной информационной системы предпринимаются в Российской Федерации на основе формирования отраслевой информационно-справочной подсистемы Министерства здравоохранения Российской Федерации по медицинскому туризму. Однако отраслевой характер разработки, вероятнее всего, не позволит охватить весь спектр услуг, которые предоставляет современный туризм, что значительно сократит потенциальную клиентскую базу и функционал такой системы.

Более рациональным здесь представляется создание функционально более широкой информационной системы, построенной на концепции «многомерного» хранилища данных в виде «куба», основными «измерениями» (dimensions) которого являются:

- региональная (иерархическая) структура туристического пространства России;
- виды туристической деятельности;

- типы (или способы) организации туризма;
- направления туристических потоков;
- нормативные затраты на предоставляемые туристические услуги;
- рыночные затраты на туристические услуги;
- основные игроки рынка туристических услуг;
- время (год, квартал, месяц).

Очевидно, что разработка такого многофункционального информационного пространства должно основываться на современных технологиях «комплексных аналитических систем» (Complete Analysis System). В этом направлении работают ведущие фирмы-производители программного обеспечения (IBM Cognos, MicroSoft Analysis Services, SAP BI и др.), которые предлагают продукты, позволяющие строить эффективные схемы моделирования, анализа и отображения данных. Эти продукты получили название Business Intelligent (BI), их объединяют схожие концептуальные схемы построения информационного обеспечения, которые отличны от известных, традиционных схем моделирования предыдущего поколения.

Большинство известных компьютерных моделей строятся по «монолитному» принципу, соединяя в одном программном блоке функции выделения, трансформации и загрузки исходных данных (так называемые ETL технологии – Extract, Transform and Load) [4], функции моделирования (Model) и функции просмотра результатов (Browser) (рис. 1). Источниками данных в таких конструкциях, как правило, являются файлы со специальной структурой, а выходные данные представляются в жестко форматированных отчетах табличного вида или в графиках.

Развитие функционала аналитики привело к формированию более сложной структуры аналитической системы класса BI, ориентированной на более глубокий, детальный анализ экономических задач и бизнес-процессов.[3] Концептуальная схема моделей класса BI существенно отличается от традиционной схемы моделирования прежде всего тем, что центральной частью BI-модели является хранилище данных (Data Ware House), которое построено

на модели многомерного «куба» (рис. 2). Кроме этого, схемы процесса ETL в этой конструкции также выделяются в отдельную компоненту, универсальную по отношению к широкому классу источников данных.

Конструкция BI-модели состоит из следующих блоков (рис. 2):

- блок исходных данных, которые могут поставляться из различных источников (текстовые файлы, табличные данные, реляционные базы данных и т.п.);
- блок хранилища данных (Ware House), в котором находятся анализируемые данные в форме «многомерного куба»;
- блок инструментов отображения, который позволяет представлять результаты моделирования в форме, удобной для восприятия с помощью OLAP-инструментов [4]. Например, в виде удобных отчетов, графиков или в форме специальных визуальных индикаторов.

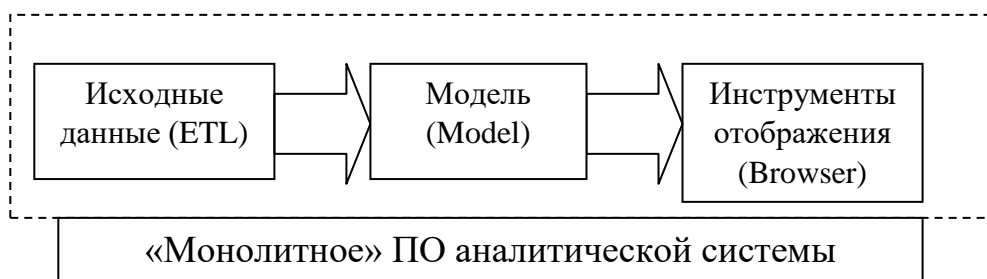


Рисунок 1 – Традиционная структура компьютерной модели («монолитная»)

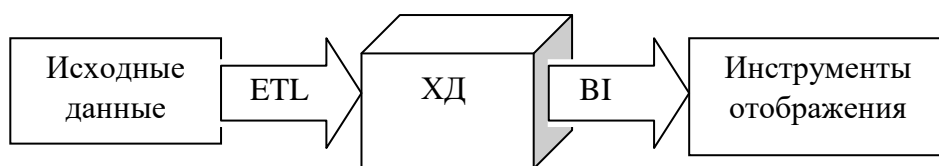


Рисунок 2 – «Классическая» структура модели BI (с хранилищем)

Процессы преобразования исходных данных в «куб» или программируются, или описываются с помощью специальных графических инструментов, позволяющих генерировать программы ETL автоматически. На выходе модели обычно используется универсальный инструмент для просмотра результатов в различных «разрезах» (OLAP-browser). Под «кубом» здесь

понимается многомерная информационная конструкция, которая состоит из таблицы «фактов» (Measure table or Fact table) и таблиц «измерений» (Dimensions). Одним из вариантов реализации концепции многомерного куба в реляционной терминологии является схема базы данных типа «звезда» (или «снежинка»), в которой центральная таблица «фактов» связана с таблицами «измерений».

Вариант реляционной модели куба для представления аналитических результатов в модели туристической отрасли РФ показан на рис. 3.

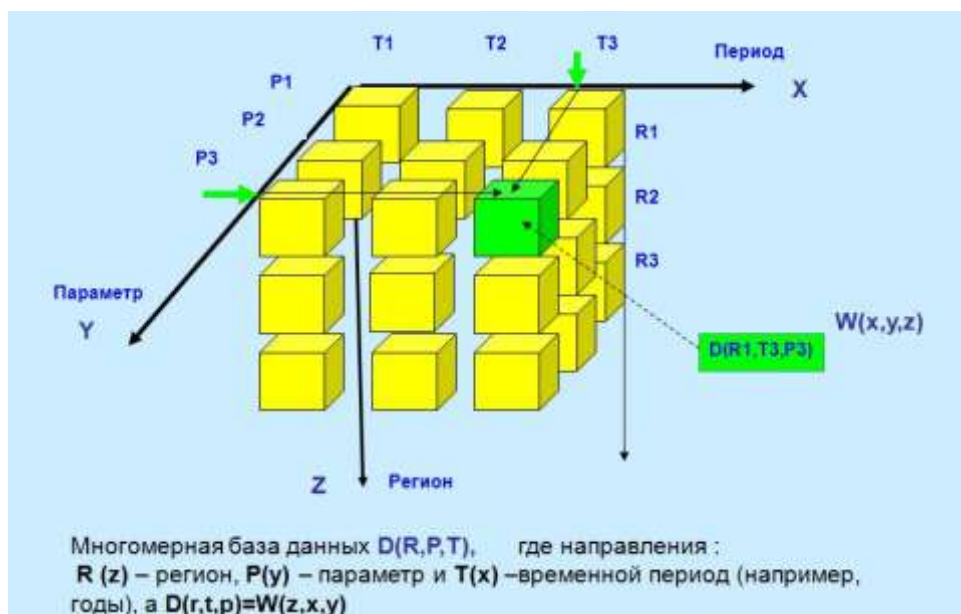


Рисунок 3 – Реляционная модель куба в модели туристической отрасли РФ

Типизация схемы хранилища данных по типу «звезда», позволяет использовать стандартные инструменты доступа к данным с помощью языка SQL или языка MDX.

Такой подход к конструированию аналитических систем значительно ускоряет сам процесс создания модельной среды, а также повышает эффективность и увеличивает функционал аналитиков.

Следующим шагом в развитии технологии построения сложных аналитических инструментов является объединение «традиционной» схемы моделирования с «классической структурой» BI.

Добавим в «классическую» схему BI блок моделирования (Model), который позволяет провести необходимые расчеты над исходными данными

(например, модели линейного программирования, прогнозные модели, модель «затраты-выпуск» и т.п.). В этом случае мы получим структуру программного комплекса (рис. 4), который позволяет не только строить модели и прогнозы экономических объектов, но и выполнять аналитические процедуры над массивом результатов моделирования.

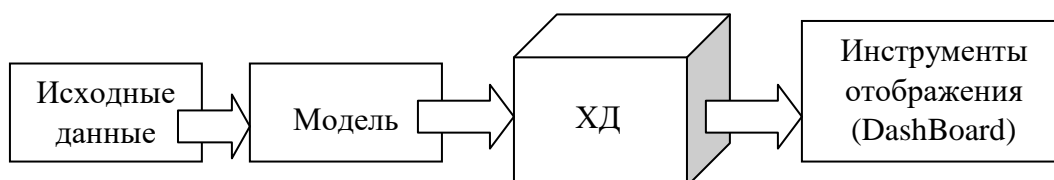


Рисунок 4 – Концептуальная схема систем класса Complete Analysis

Выводы: рассматривая различные подходы к разработке единого информационного пространства туристического бизнеса, следует остановиться на прогрессивных технологиях, объединенных в концепцию, которая сочетает в себе подходы как «традиционных» информационных моделей, так и «классического» ВІ (т.е. Complete Analysis System (CAS)).

## ЛИТЕРАТУРА

1. Официальный сайт UNWTO. URL: <http://www.e-unwto.org/doi/pdf/10.18111/9789284416899> (дата обращения: 01.12.2016).
2. Официальный сайт Deloitte Touche Tohmatsu Limited. URL: <https://www2.deloitte.com/content/dam/Deloitte/global/Documents/Life-Sciences-Health-Care/gx-lshc-2016-health-care-outlook.pdf> (дата обращения: 01.12.2016).
3. Герасимов Н.А. Один подход к разработке хранилища данных для модели мировой экономики – Вестник Российского университета имени Г.В. Плеханова, ФГБОУ ВО «им. Г.В. Плеханова», №2 (86), 2016, с. 152-157.
4. Паклин Н.Б., Орешков В.И. Бизнес-аналитика: от данных к знаниям (+CD). СПб.: Питер, 2014. 709 с.