



Е. В. Колесниченко, С. Е. Абдыкеров

ПРИМЕНЕНИЕ ТЕХНОЛОГИЙ BUSINESS INTELLIGENCE ПРИ СТРАТЕГИЧЕСКОМ АНАЛИЗЕ ДЕЯТЕЛЬНОСТИ ТЕХНИЧЕСКОГО УНИВЕРСИТЕТА



E. V. Kolesnichenko, S. E. Abdykherov

Adoption of Business Intelligence technologies in strategic analysis of operating process in the technical university

Strategic management of the university is impossible without registering process of both internal and external factors. An effective means of infotainment of the Decision Support System are the Business Intelligence technologies. The article undercovers experience of practical application of the BI technologies in Moscow Aviation Institute.

К л ю ч е в ы е с л о в а: стратегический анализ, система поддержки принятия решений, интеллектуальный анализ данных, интеллектуальный анализ текстов, оперативный анализ данных, контент-мониторинг.

К e y w o r d s: strategic analysis, business intelligence, data mining, text mining, OLAP, content monitoring.

Возрастание конкуренции на рынке образовательных услуг, неблагоприятная демографическая ситуация, ограниченность финансирования и ряд других факторов внешней среды заставляют руководство большинства российских университетов пересматривать привычные методы деятельности, сделав акцент на стратегическом развитии вузовского образования и науки, учитывающем долговременные перспективы [4, 5]. Отчетливо осознавая эту тенденцию, новое руководство Московского авиационного института (МАИ) не стоит в стороне от назревших проблем и целенаправленно проводит в жизнь стратегическую линию укрепления своей ведущей роли как головного вуза аэрокосмической отрасли. Целью данной статьи является рассмотрение некоторых подходов, принятых ректоратом МАИ для обеспечения идеи стратегического лидерства, которые, как мы надеемся, будут

полезны вузовскому сообществу в практической деятельности.

Как показывает практика [1, 3], для того чтобы внедряемая система стратегического менеджмента была действительно эффективной, а не просто декларативной, вузу необходимо пройти три основных этапа, тесно связанных друг с другом и образующих единый контур управления:

— стратегический анализ, целью которого является обеспечение руководства университета полной, оперативной и достоверной информацией, необходимой для принятия взвешенных стратегических решений;

— формализация стратегии, заключающаяся в формулировке стратегических целей и задач вуза в соответствии с его миссией, ценностями и видением;

— реализация стратегии, т. е. разработка кон-

кретных стратегических инициатив и планов, как корпоративных, так и личных.

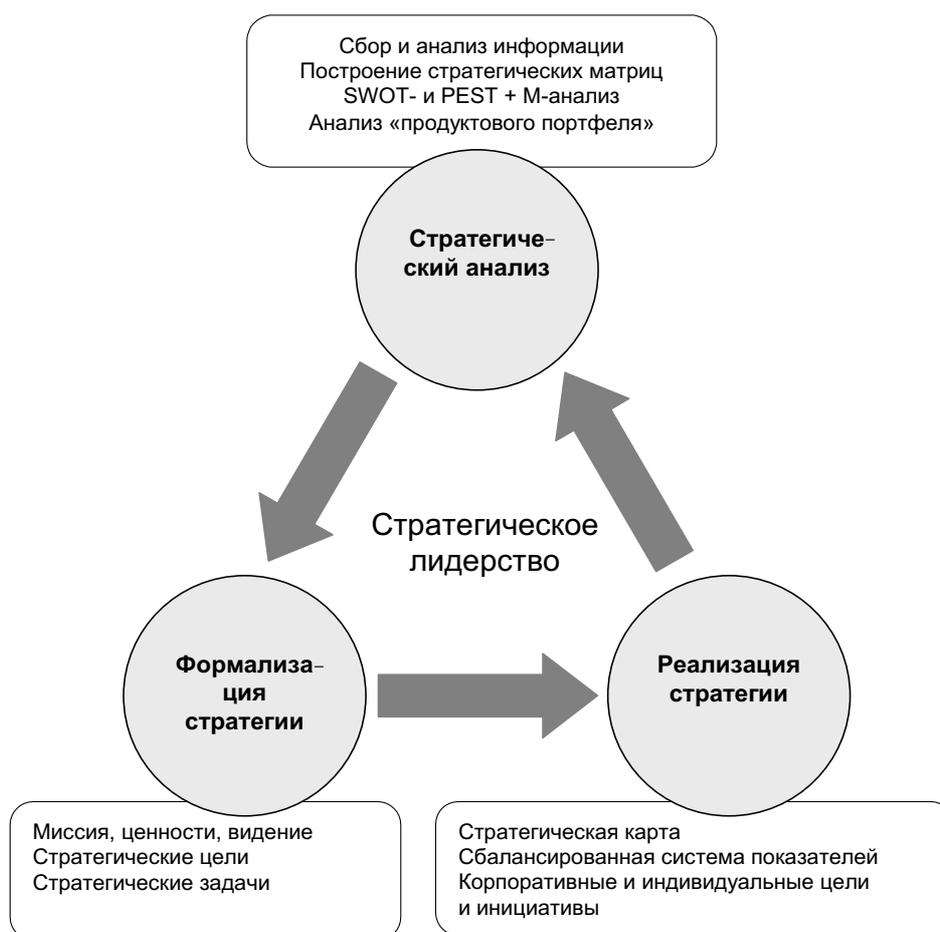
На рисунке представлены этапы стратегического управления и их основные элементы. Данная схема является достаточно условной, поскольку зачастую некоторые этапы, например стратегический анализ и формализация стратегии, проводятся параллельно. Необходимо отдавать себе отчет в том, что внедрение стратегического управления является многошаговым итеративным процессом, при котором для достижения оптимальных решений необходимо последовательно неоднократно повторять все этапы от начала до конца. Каждая такая итерация дает новое понимание ситуации, позволяет уточнить и скорректировать ранее принятые решения.

Учитывая ограниченность объема статьи, ниже рассматриваются только некоторые аспекты первого из отмеченных этапов, от полноты и достоверности которого во многом зависят результаты всей работы. Его основная задача — сбор информации об основных направлениях де-

ятельности вуза, ее критический анализ с целью получения объективной оценки состояния дел в различных подразделениях и университете в целом, определение текущих приоритетов и долгосрочных перспектив развития.

Одной из ключевых проблем этапа стратегического анализа является информационная составляющая, т. е. консолидация разнородной по своему составу и представлению информации в единой базе данных, возможность оперативного ее извлечения в детализированном или агрегированном виде для последующего проведения бизнес-анализа. Как свидетельствует мировой опыт, такая многофункциональность систем поддержки принятия управленческих решений может быть достигнута в рамках концепции Business Intelligence (BI). Впервые термин «Business Intelligence» был введен в обращение аналитиками компании Gartner в конце 1980-х гг. и в настоящее время трактуется как:

— процесс превращения данных в информацию и знания о бизнесе для поддержки принятия улучшенных и неформальных решений;



Этапы стратегического управления и их основные элементы

— информационные технологии (методы и средства) сбора данных, консолидации информации и обеспечения доступа пользователей к знаниям;

— знания о бизнесе, добытые в результате углубленного анализа детальных данных и консолидированной информации.

Актуальность и практическую ценность этих инновационных подходов в области стратегического управления подтверждают исследования «META Group», согласно которым до 95 % ведущих мировых компаний активно применяют технологии BI, добиваясь преимуществ в конкурентной борьбе и получая значительную отдачу от своих инвестиций. По данным «International Data Corporation», в течение трех лет они в среднем получили 400-процентный возврат своих инвестиций в BI-системы.

Концепция Business Intelligence базируется на трех китах [2, 8]:

- хранилище данных (data warehouse);
- средства оперативного анализа данных (OLAP);
- средства интеллектуального анализа данных (data mining).

Одним из важнейших компонентов BI-системы считается хранилище данных (ХД) — специализированная база для поддержки процесса анализа данных. Базовыми требованиями к ХД являются: ориентация на предметную область, интегрированность и внутренняя непротиворечивость, привязка ко времени, неизменяемость, поддержка высокой скорости получения данных из хранилища, полнота и достоверность хранимых данных, поддержка качественного процесса пополнения данных.

Кроме хранилища данных, неизменным атрибутом BI-системы являются средства интеллектуального анализа данных и текстов, которые возникли на стыке прикладной статистики, теории баз данных, методов искусственного интеллекта, теории информации и ряда других дисциплин и обеспечивают процесс поддержки принятия решений, основанный на поиске в данных (числовых или текстовых) скрытых закономерностей. При этом накопленные сведения автоматически обобщаются до информации, которая может быть охарактеризована как новые, ранее неизвестные, нетривиальные, практически полезные и доступные для интерпретации знания.

Третья составная часть BI-систем — средства оперативного анализа данных, или OLAP (on-line analytical processing) — технология комплексного динамического синтеза, анализа и кон-

солидации больших объемов многомерных данных. Системы OLAP-анализа представляют данные в виде многомерных наборов, называемых многомерными кубами, оси которых содержат основные атрибуты анализируемого бизнес-процесса, а ячейки — зависящие от них агрегированные данные, количественно его характеризующие. Применение средств OLAP, агрегирующих информацию за годы и позволяющих быстро просматривать данные в любом желаемом разрезе, значительно упрощает решение стратегических задач.

При разработке концепции единой информационной системы (ЕИС) МАИ за основу были приняты описанные выше подходы к построению BI-систем. Важнейшим этапом создания такой системы является разработка информационной модели, отражающей существующие инфокоммуникационные процессы в организации, поскольку тщательность проработки этого этапа оказывает существенное влияние на дальнейшие работы по развитию системы, гибкости ее в настройке. В основе информационной модели должны лежать информационные потоки, циркулирующие внутри подразделений предприятия, вместе с этим согласно концепции BI она также должна учитывать нахождение организации в информационном поле, не ограниченном только собственными рамками предприятия.

Применительно к ЕИС МАИ первичными являются данные, находящиеся в обращении следующих подразделений:

- факультеты;
- отдел аспирантуры и докторантуры;
- бюро приказов;
- отдел кадров и т. д.

Кроме отмеченных выше внутренних коммуникаций, ЕИС должна поддерживать также и внешние информационные связи, в частности взаимодействие с министерством, РФФИ, налоговыми органами, заказчиками и другими организациями. Системный анализ таких инфокоммуникационных процессов позволяет оптимальным образом спроектировать структуру ЕИС и выбрать архитектуру хранилища данных.

Одним из основных элементов ЕИС является система анализа данных, построенная на базе отечественной аналитической платформы Deductor [7] и обладающая развитыми средствами Data Mining. Ее применение позволяет в онлайн-режиме решать текущие задачи стратегического анализа, используя накопленную в едином хранилище данных информацию. Здесь хотелось бы остановиться на одном аспекте прак-

тического использования ВІ-систем, который особенно важен для вузов при разработке стратегии (в первую очередь при проведении SWOT- или PEST + М-анализа), — экспертизе перспективности проводимых научно-исследовательских работ.

Задача такой экспертизы — получение объективной оценки востребованности на текущий момент того или иного научного направления деятельности вуза. К сожалению, как показывает практика, традиционный подход, когда учитываются только мнения членов экспертного совета, проректоров, деканов, руководителей НИО факультетов и других уважаемых людей университета, во многом носит субъективный характер вследствие конъюнктурных соображений, личных симпатий, консерватизма научных взглядов и ряда других известных факторов. Поэтому для принятия взвешенных стратегических решений кроме учета таких мнений необходима дополнительная информация, которая бы давала объективную картину тенденций развития соответствующих направлений мировой науки и техники.

Исходя из этих соображений, на основе ВІ-технологий была разработана методика контент-мониторинга информационных ресурсов Интернета с целью анализа тенденций развития фундаментальных и прикладных научных направлений университетской науки. Идея такого контент-мониторинга как инструмента анализа не является оригинальной и применительно к указанным выше задачам заключается в том, чтобы с помощью набора ключевых слов, отражающих специфику конкретного научного направления, проследить временную динамику изменения интереса к нему в различных научных информационных источниках. Достоинствами такой методики являются максимальная объективность оценок, однозначность интерпретации результатов, оперативность получаемых оценок, высокая степень автоматизации процесса, возможность среднесрочного прогнозирования.

При реализации методики решались следующие основные задачи:

- определение перечня научных баз данных в соответствии с их доступностью (платные / бесплатные ресурсы) и информативностью;
- градация информационных источников по степени их достоверности и авторитетности в научной среде;
- формирование поискового запроса, отвечающего условиям смысловой релевантности;
- количественная оценка тенденций развития процессов и их классификация.

Среди научных баз были рассмотрены как

коммерческие (STN International и Scopus), так и бесплатные (научная поисковая система Scirus, база ISI Web of Knowledge, научная поисковая система Scholar google, научная поисковая система Scienceresearch, база данных Стэнфордского университета High Wire Press, научная поисковая система High Beam Research). Практическая работа с бесплатными базами показала, что наиболее полной и удобной для работы (возможность задания поиска по годам, градация источников информации по категориям и т. д.) является научная поисковая система Scirus. Что касается платных баз, то, конечно, их информативность и качество сервиса оказались значительно выше, однако результаты работы показали, что *качественная картина* временной динамики изменения числа публикаций в Scirus и Scopus (годовая подписка от 30 тыс. евро) примерно идентична.

Любая информационно-поисковая система в ответ на запрос, сформированный определенным образом из ключевых слов, выдает перечень найденных документов, которые необходимо отсортировать в соответствии с их достоверностью. В Scirus решение этой задачи облегчается тем, что найденные источники изначально группируются в три класса:

- Journal sources, включающий в себя книги, журналы и материалы научно-технических конференций;
- Preferred Web, основу которого составляют патентные базы данных, препринты, диссертации, технические доклады и университетские библиотеки;
- Other Web — это университетские сайты, домашние страницы ученых, новостные сайты и сайты различных организаций.

Результаты информационных поисков показали, что для анализа научных направлений университетов наиболее приемлемыми являются два первых класса (*Journal sources* и *Preferred Web*), так как они содержат наиболее авторитетные источники научно-технической информации. Как правило, тенденции (хронология числа публикаций) в обоих этих классах идентичны. Что касается третьего класса, то в последние несколько лет в связи развитием семантического Интернета здесь наблюдается лавинообразный рост числа публикаций, в своей массе весьма спорных и, как правило, слабо коррелированных с тематикой запросов.

Третья задача контент-мониторинга заключалась в формировании поискового запроса и отборе документов в соответствии с их смысловой

релевантностью — задача, решаемая на уровне технологий Text Mining [6]. В существующих научных базах данных таких алгоритмов нет, так как смысловой отбор документов полностью отнесен к компетенции пользователей. По существу, при проведении поиска по множеству ключевых слов пользователь поставлен перед дилеммой:

— вводить их все с помощью оператора OR и получить при этом с большой вероятностью массу документов, имеющих весьма далекое отношение к тематике запроса (как показывает практика, их число достигает 60–70 %);

— вводить их все с помощью оператора AND и получить в 90 % случаев ответ «По вашему запросу источников в базе найти не удалось».

Для того чтобы получить репрезентативную выборку документов, которые бы в максимальной степени отвечали требованиям смысловой релевантности, применительно к научным базам данных был разработан оригинальный многошаговый алгоритм формирования поисковых запросов и обработки списков документов. Здесь мы не будем касаться деталей данного алгоритма, отметим только, что он позволяет выделить те документы, в которых встречается не менее N ($N = 2, 3, \dots$) ключевых слов, при этом чем их больше в документе, тем больше его удельный вес.

И, наконец, на заключительном этапе контент-мониторинга осуществлялось агрегирование полученных в результате информационного поиска результатов. На этом этапе в полной мере использовались возможности когнитивных технологий Data Mining, реализованных в платформе Deductor, по классификации процесса, т. е. отнесению его к одному из нескольких предопределенных классов. Применительно к научным направлениям на основе экспертных оценок было выделено пять таких классов (в принципе, их число может быть произвольным):

— процесс с явной доминирующей тенденцией роста, соответствующий стремительно развивающейся научной тематике;

— процесс с положительной динамикой роста, отражающий постоянно возрастающий интерес к тематике;

— процесс равновесного состояния — стадия зстоя и стагнации;

— процесс с отрицательной динамикой спада (обратный классу 2);

— процесс с явной доминирующей тенденцией спада (обратный классу 1).

Таким образом, имея в качестве исходной информации множество ключевых слов, отражающих предметную область конкретной научной тематики, в качестве агрегированного выходного результата контент-мониторинга мы получаем номер класса, к которому она относится, т. е. интегрированную характеристику данной научной тематики. При необходимости эта информация может быть детализована: таблицы и статистики по группам источников и годам, временные диаграммы с трендами и т. д.

Конечно, при принятии решений надо очень осторожно подходить к использованию полученных результатов и учитывать специфику конкретной тематики, ведь далеко не всегда уменьшение числа открытых публикаций свидетельствует о снижении интереса к ней. Но в любом случае, имея на руках такие данные, руководство вуза может более взвешенно подходить к финансированию перспективных направлений своей научной деятельности, учитывая реально складывающиеся мировые тенденции.

В заключение хотелось бы отметить, что даже на одном рассмотренном выше примере видно, насколько эффективно практическое применение технологий Business Intelligence в качестве инструмента анализа и прогнозирования, эти технологии позволяют освободить принимающих решения лиц от рутинной работы, копания в кипе отчетов и дают возможность видеть картину в целом.

1. Аакер Д. Стратегическое рыночное управление. 7-е изд. / пер. с англ. ; под ред. С. Г. Божук. СПб. : Питер, 2007. 496 с.

2. Базы данных. Интеллектуальная обработка информации / В. В. Корнеев, А. Ф. Гареев, С. В. Васютин, В. В. Райх. М. : Нолидж, 2003. 400 с.

3. Виханский О. С., Наумов А. И. Менеджмент. 3-е изд. М. : Экономистъ, 2003. 528 с.

4. Клюев А. К., Корунов С. М. Стратегии вузовского развития // Университетское управление: практика и анализ. 2003. № 3. С. 43–50.

5. Князев Е. А. Об университетах и их стратегиях // Университетское управление: практика и анализ. 2005. № 4. С. 9–17.

6. Ландэ Д. В. Поиск знаний в Internet. М. : Изд. дом «Диалектика-Вильямс», 2005. 272 с.

7. Материалы сайта компании Base Group. URL: <http://www.basegroup.ru>

8. McKnight W., Humphrey S. Building Business Intelligence: Rafting Into the Business Intelligence Future, Part 1 // DM Review Magazine. 2004. Oct.