

ПРЕДСТАВЛЕНИЕ БИЗНЕС-ПРОЦЕССОВ ОРГАНИЗАЦИИ НА ОСНОВЕ ТЕХНОЛОГИИ TEXT MINING

Рассматриваются вопросы построения бизнес-процессов на основе интеллектуального анализа текстов документов, описывающих процессы, которые происходят в организации, - должностных инструкций, договоров, ТУ, регламента работ и т.п.

1. Актуальность проблемы и постановка задачи исследования

В настоящее время в отечественных организациях создается множество отличающихся друг от друга, не связанных между собой, недостаточно проработанных документов, которые фактически предназначены для решения задачи регламентации реализующихся в организации бизнес-процессов (БП) в целях эффективного управления данной организацией. Для организации и ее собственников существующее положение дел означает неэффективное расходование финансовых, материальных, человеческих, временных ресурсов, потерю конкурентных преимуществ на рынке.

В связи с этим возникает задача создания, на основе существующей документации и неформальных знаний сотрудников организации, единой системы описания и управления БП, которая позволяла бы решать основные задачи поддержки принятия решений при управлении организацией и ее развитии на основании процессного подхода и обеспечивала бы достижение следующих целей:

1. Непрерывное улучшение результатов деятельности организации на основе планирования и учета соответствующих показателей.
2. Удовлетворение потребностей инвесторов, учредителей, потребителей, поставщиков, персонала предприятия.

Система описания БП должна решать задачи регламентации:

- порядка планирования целей и деятельности;
- взаимодействия между процессами и подразделениями организации;
- ответственности и полномочий должностных лиц, в том числе владельцев процессов;
- порядка работы и действий в нештатных ситуациях;
- порядка и форм отчетности перед высшим руководством;
- системы показателей, характеризующих результативность и эффективность деятельности организации в целом и ее процессов;
- порядка рассмотрения результатов деятельности и принятия управленческих решений по устранению отклонений и достижению плановых показателей.

Для решения задач, стоящих перед системой описания и управления БП, необходимо сформировать показатели эффективности управления, которые должны включать в себя:

- показатели результата деятельности отдельных процессов и организации в целом (достижение запланированных результатов - по объему, качеству, номенклатуре и срокам);
- показатели эффективности деятельности отдельных процессов и организации в целом (отношение полученных результатов к затратам времени, финансовых и других ресурсов);
- параметры продукции, производимой процессами организации;
- показатели удовлетворенности клиентов результатами работы.

Рассмотренные показатели могут использоваться как для оценки своей работы исполнителем БП, так и для отчетности перед руководством либо контролирующими организациями.

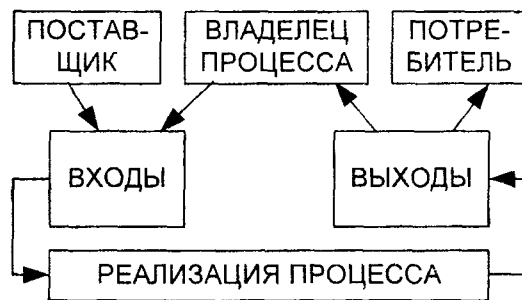
Следует отметить, что необходимым условием для организации процессного управления является наличие на предприятии информационной системы.

2. Анализ процессного подхода к управлению организацией

Процессный подход к управлению организацией основан на формализованном описании ее функционирования в виде совокупности взаимодействующих БП.

В соответствии со стандартом ISO 9000:2000, «Процесс – совокупность взаимосвязанных или взаимодействующих видов деятельности, преобразующая «входы» в «выходы». Соответственно, бизнес-процесс можно определить как устойчивую, целенаправленную совокупность взаимосвязанных видов деятельности (последовательности работ), которая по определенной технологии преобразует входы в выходы, представляющие ценность для потребителя [1].

Обобщенная схема БП представлена на рисунке.



Обобщенная схема бизнес-процесса

Входы БП представляют собой ресурсы, обеспечиваемые как внешним поставщиком, так и владельцем процесса.

Ресурсы – это материальные и нематериальные компоненты, необходимые для достижения целей процесса и находящиеся в распоряжении владельца БП, например, технология производства, оборудование, персонал, инфраструктура, программное и аппаратное обеспечение.

Владелец БП – это должностное лицо, которое имеет в своем распоряжении необходимые для выполнения БП ресурсы, управляет его ходом и несет ответственность за результаты и эффективность БП. Для эффективного управления БП владелец, помимо соответствующих ресурсов, должен иметь документированную технологию реализации процесса, а также систему сбора информации о текущих параметрах процесса, характеристиках продукта и их соответствии требованиям потребителя.

Выход БП – это результат (продукт, услуга) выполнения БП, а также информация о его выполнении, которая передается владельцу БП, например, текущие параметры процесса, характеристики продукта, степень удовлетворенности потребителя (заказчика).

Потребители подразделяются на внешних и внутренних. Внутренние находятся в той же организации, в которой реализуется БП, и в ходе своей деятельности используют его результаты.

Внешние потребители находятся за пределами предприятия и используют результат БП.

БП подразделяются на основные и вспомогательные. Основные процессы создают прямую ценность для потребителя. Их результатом обычно являются продукты (услуги).

Вспомогательные (поддерживающие) процессы необходимы для функционирования других процессов и создают косвенную ценность для потребителя.

БП целесообразно разбить на бизнес-процедуры. Бизнес-процедура определяется следующим набором параметров: результат; исполнитель; действие; основания для осуществления действия; входные данные.

Результат бизнес-процедуры является промежуточным результатом БП и, одновременно, входными данными в следующую бизнес-процедуру (сырьем для нее). Результат бизнес-процедуры должен учитываться, контролироваться и управляться с точки зрения конечного продукта БП. Для результата

процедуры необходимо сформировать набор требований, описываемых в количественных и качественных показателях и определяющих его качество и условия его достижения. Источником для такого набора требований являются нормативные, административные и другие документы, а также информация от должностных лиц. Сформированные требования и показатели должны быть доступны для контроля.

Исполнителем бизнес-процедуры обычно является должностное лицо или подразделение организации.

Действие бизнес-процедуры осуществляется исполнителем над входными данными для получения результата.

Основания для осуществления действия наделяют исполнителя соответствующими полномочиями и содержатся в административных и распорядительных документах и указаниях руководства.

Формализация БП является необходимым условием для реализации процессного подхода к управлению организацией и включает в себя выполнение работ:

- по описанию БП;
- по регламентации БП;
- по проведению аудита БП.

Описание БП заключается в создании документации, содержащей информацию о текущем состоянии БП и обеспечивающей решение следующих задач:

- анализ и оптимизация БП по заданным показателям эффективности;
- формирование эффективной системы управления, основанной на процессном подходе;
- разработка нормативных документов, регламентирующих технологию выполнения БП (либо его отдельных элементов);
- подготовка к созданию и внедрению информационно-управляющих систем.

Подходы к описанию основных и вспомогательных БП несколько различаются. Основные процессы определяют пути создания продукции на предприятии и, соответственно, отражают жизненный цикл продукции. К таким процессам, в соответствии с рекомендациями Международной Бенчмаркинговой Палаты (International Benchmarking Clearinghouse), относятся:

- маркетинг рынка и пожеланий заказчиков;
- разработка стратегии;
- разработка продукции (услуг);
- организация продаж;
- производство и поставка продукции;
- организация сервиса (для организаций, ориентированных на сервисное обслуживание);
- обслуживание заказчика и выписка счета-фактуры;
- управление человеческими ресурсами;
- управление информационными ресурсами;
- управление финансовыми и материальными ресурсами;
- экологический мониторинг;
- управление внешними связями;
- управление улучшениями и изменениями.

Следует отметить, что на практике при выделении основных процессов используется известное из курса психологии ограничение 7 объектов: человек без затруднений одновременно может обрабатывать информацию, получаемую не более чем от 7 (иногда от 9) объектов либо процессов.

Вспомогательные процессы обеспечивают нормальное функционирование основных процессов и обычно включают в себя:

- управление персоналом;
- управление финансами;

- управление документооборотом предприятия;
- управление внешними связями;
- управление развитием.

При регламентации БП разрабатываются новые (либо дорабатываются существующие) нормативные документы, регламентирующие выполнение операций, входящих в состав БП. Такие документы должны включать в себя показатели БП, методологию их измерения, порядок отчетности о состоянии процесса и т.п.

При проведении аудита БП проверяется соответствие существующего БП требованиям, установленным в нормативных и организационных документах, а также оценка эффективности БП по заданным показателям. Результатом аудита БП является отчет, который содержит описание текущего процесса и предложения по переходу к новой, более эффективной его организации.

3. Представление бизнес-процессов организации на основе технологии *text mining*

Недостатком процессного подхода в чистом виде является то, что он входит в противоречие с существующей функциональной структурой предприятия и требует матричного способа управления. Матричное управление означает, что процесс, охватывающий несколько подразделений, имеет владельца, который не является их руководителем. Следовательно, размываются ресурсы и ответственность за результаты выполнения БП, могут возникать конфликты между владельцем БП и руководителями подразделений. Как показывает зарубежный опыт, переход ведущих фирм к матричному управлению занимает несколько лет.

В связи с указанным недостатком целесообразным представляется эволюционное развитие процессного подхода на предприятии с учетом существующей функциональной структуры и ее плавным изменением. Такому решению проблемы способствуют следующие факторы:

- функции, реализуемые подразделениями, во многих случаях можно рассматривать в качестве подпроцессов, либо бизнес-задач, являющихся элементами БП;
- одним из важнейших условий при описании БП является наличие всей необходимой документации, которая формируется функциональными подразделениями;
- предприятие, развиваясь в условиях рынка, изменяет свою функциональную структуру для ее максимального соответствия тактическим и стратегическим целям.

Таким образом, при реализации процессного подхода к управлению организацией возникают две проблемы:

- анализ существующей структуры и прогнозирование изменений в структуре организации в целях ее максимального соответствия стоящим перед организацией задачам;
- автоматизация описания БП на предприятии на основе использования существующих нормативных, справочных и организационных документов.

Первая проблема характеризуется:

- влиянием значительного числа случайных (и противоречивых) факторов;
- регулярностью выполнения (в изменяющихся рыночных условиях требуется постоянная адаптация организации с учетом имеющегося опыта, появления новых заказчиков, выпуска новой продукции конкурентами и т.п.).

Для ее решения целесообразно использовать методы интеллектуального анализа данных и методы эволюционного моделирования.

Методы интеллектуального анализа данных (*data mining*) предназначены для нахождения правил в базах данных организации, которые, в общем случае, содержат в неявном виде информацию о закономерностях происшедших с организацией событий.

Методы эволюционного моделирования основаны на использовании закономерностей естественного отбора для моделирования развития технических, экономических, организационных систем, решения комбинаторных задач и т.п.

Вторая проблема относится к области искусственного интеллекта и заключается в преобразовании смысловых взаимоотношений между элементами текста в структурное описание, которое служило бы основой системы поддержки принятия решений при формализации элементов БП.

Для ее решения целесообразно использовать методы интеллектуального анализа текста (text mining).

Text mining можно определить как процесс анализа текста, направленный на извлечение полезной для специфических целей информации [2].

Потребность в интеллектуальном анализе текста возникла в связи с тем, что в настоящее время традиционные методы обработки данных не соответствуют быстро растущим объемам информации (в частности, в электронной форме).

Возможность интеллектуального анализа текста определяется тем, что для извлечения из него полезной информации совершенно не обязательно понимать этот текст. Текстовые документы часто скрывают ценные структурные данные. Например, совокупность газетных статей может содержать информацию относительно местоположения офисов множества организаций. Интернет содержит миллионы страниц, данные на которых можно преобразовать в структурную форму.

В области интеллектуального анализа данных решаются следующие основные задачи:

- нормализация текста;
- автоматическая категоризация и выделение ключевых слов в документе;
- сегментация текста по содержанию;
- идентификация разделов и отслеживание временных параметров документов;
- идентификация языка текстового документа;
- определение авторства текстовых документов;
- идентификация копий документов;
- визуализация документов;
- создание и анализ пользовательских профилей;
- сжатие текстовых документов.

Нормализация слов в тексте выполняется в 2 этапа:

- удаление служебных слов;
- нормализация слов в тексте.

В текстах на естественных языках существует целый ряд слов, которые, с лингвистической точки зрения, несут минимум информации. Очевидно, что такие слова зависят от языка. Например, для английского языка это слова: A, ABOUT, ABOVE, ACROSS, AFTER, AGAIN, AGAINST, ALL, ALMOST, ALONE, ALONG, ALREADY, ALSO, для русского: О (чем-то), НУ, ВОТ, СКОРА, ВСЕ, ПОЧТИ, УЖЕ, ТАКЖЕ. При выполнении нормализации такие слова удаляются из текста.

Следует отметить, что серьезную проблему при анализе текста вызывают различные формы одного и того же слова, имеющие разное произношение, но близкое значение. Для решения этой проблемы все слова в тексте приводятся к единой нормальной форме. Этот процесс называется стеммингом (stemming).

Автоматическая категоризация документов заключается в том, чтобы на основании набора распределенных по категориям документов автоматически присваивать вновь поступающим документам одну или несколько релевантных категорий. При этом категории могут быть как структурированными, так и неструктурированными. Ключевые слова в документах определяются аналогичным образом. Важным требованием к категоризации документов является независимость от их языка.

Один из наиболее распространенных подходов к категоризации документов включает в себя следующие основные шаги:

Шаг 1. Создание иерархии категорий документов.

Шаг 2. Заполнение иерархии категорий документами вручную.

Шаг 3. Создание для каждой категории краткого описания и набора ключевых слов.

Шаг 4. Построение модели категоризации на основании полученного в шагах 1-3 массива данных с использованием методов машинного обучения.

Шаг 5. Автоматическая категоризация новых документов и присвоение ключевых слов на основании построенной модели категоризации.

Задача сегментации текста по его содержанию состоит в том, чтобы разделить текст, не имеющий заданной структуры, на фрагменты, имеющие похожее содержание. Основная область применения приложений сегментации текста заключается в поиске заданных тем в больших текстовых базах данных, а также в новостях.

Сегментация текста включает в себя следующие шаги:

- выделение предложений в тексте;
- выделение слов и фраз в каждом предложении;
- вычисление степени близости между парами предложений;
- поиск сегментов текста на основании степени близости между парами предложений, при этом степень близости между предложениями, относящимися к одному сегменту, должна быть максимальной, а к различным сегментам – минимальной.

Идентификация разделов документа и отслеживание временных параметров документов заключается в следующем. На основании временных параметров документов, например, новостей, необходимо идентифицировать документы, которые содержат новые темы, а также определить документы, принадлежащие к существующим темам.

Задача идентификации языка текстового документа определяется тем, что системы анализа текста в общем случае используют ряд методов, зависящих от языка обрабатываемых документов.

Обобщенный алгоритм идентификации языка состоит из следующих шагов:

- построение таблицы частотных характеристик пар и троек букв для каждого языка;
- идентификация языка документа на основании таблицы частотных характеристик букв;
- в том случае, если определить язык не удалось (например, для коротких документов), устанавливаются служебные слова, которые зависят от языка.

Определение наиболее вероятных авторов для новых документов на основании существующей базы документов и их авторов является целью задачи определения авторства документов. Решение задачи основывается на том факте, что для каждого автора характерна своя частота повторения слов и фраз в тексте.

Проблема идентификации копий документов состоит в необходимости определить, что данный документ представляет собой копию (полную либо частичную) других документов, имеющихся в базе данных. Решение данной задачи основывается на индексировании частей этого документа различной длины и сравнении их с существующими документами.

Визуализация больших массивов текстовых документов состоит в преобразовании их в направленные сети и в дальнейшей визуализации этих сетей.

Составление пользовательских профилей выполняется на основе анализа текстовых баз данных, содержащих информацию о поведении пользователя, например log-файлов, cookies и т.п.

В настоящее время разработка и использование приложений интеллектуального анализа данных ведется, в основном, применительно к Интернет по следующим направлениям:

- поисковые серверы для больших текстовых баз данных;
- системы автоматической категоризации документов;
- системы построения набора ключевых слов для документов;
- интеллектуальные агенты фильтрации информации;
- интеллектуальные агенты поиска в Интернет;

- интеллектуальные документы просмотра информации в Интернет;
- интеллектуальные агенты поиска экспертов в Интернет;
- электронные библиотеки.

4. Представление бизнес-процессов на основе интеллектуального анализа документов

Интеллектуальный анализ нормативных, справочных и организационных документов позволяет, как было показано выше, получить структурные данные, содержащие знания о протекании БП в организации. Рассмотрим, в соответствии с данным подходом, вопросы автоматизации процессов обработки заявок заказчиков на примере предприятия по конструированию и производству низковольтной электроаппаратуры.

Отправной точкой таких процессов является получение документа с заказом, заявкой, договором на поставку продукции, документации, сервисное обслуживание. Документы могут приходить по разным каналам (факс, электронная почта, обычная почта), однако в большинстве случаев все они имеют близкую структуру:

- наименование и реквизиты заказчика;
- информация о пожеланиях заказчика, где указан тип товара (услуги), который ему необходим – информация о характеристиках и ценах товара, документация, разработка нового изделия, поставка партии товаров;
- наименование и количество заказываемого товара, где обычно указывается класс товаров (муфты, тормоза, выключатели) и код модели товара, либо код документации (например, технических условий);
- дополнительные требования, определяющие вид исполнения товара (тропическое, для работы при пониженных температурах);
- желаемые сроки поставки.

Документы принимает технический работник, после чего менеджер предприятия, проанализировав заказ, распределяет его по соответствующим подразделениям с назначением ответственного за обработку заказа.

В процессе распределения заявок между исполнителями учитываются следующие факторы:

- предприятие (обработка заявок и выполнение заказов для крупных предприятий, обеспечивающих долговременные заказы и высокую прибыль, выполняется в первую очередь);
- сроки поставки;
- тип товара либо услуги;
- сроки исполнения;
- стоимость.

Таким образом, обобщенный БП обработки заявок заказчиков предприятия по производству низковольтной аппаратуры включает следующие основные процедуры:

- выявление наименования предприятия в тексте документа;
- классификация предприятия;
- классификация заказов (документация, поставки);
- выявление наименования и кода заказываемого товара, а также дополнительных требований к нему;
- выявление объемов и сроков поставки;
- формирование предложений по исполнителям данного заказа.

Для детального описания и автоматизированной корректировки рассмотренного обобщенного БП необходимо выполнить следующие основные шаги:

- а) нормализацию слов в тексте – приведение всех вариантов одного и того же слова к нормальной форме;
- б) выделение и нормализацию фраз в тексте, а также расшифровку сокращений, которые обозначают:
 - заказчиков;

- сроки поставки;
- тип товара либо услуги;
- наименование изделия;
- в) составление списка нормализованных ключевых слов, что позволит однозначно определить тип, а также наименование заказываемого товара (услуги);
- г) категоризацию документов по типам товара (услуги) на основании списка ключевых слов;
- д) категоризацию документов по наименованиям заказываемого товара на основании списка ключевых слов;
- е) идентификацию временных параметров документов, что дает возможность определить срочность документа и скорректировать сроки поставки;
- ж) составление списка предлагаемых исполнителей по каждому документу на основании категории документа и базы данных исполнителей.

5. Выводы

Рассмотренный подход к описанию БП позволяет автоматизировать процесс извлечения знаний из текстовых документов и служит основой для поддержки принятия решений при описании и управлении БП.

Система описания и управления БП обеспечивает достижение следующих преимуществ:

- гибкость и прозрачность управления организацией;
- возможность оценки эффективности управления на каждом этапе производственного и управленческого процесса;
- формализация документооборота в организации и приведение его в соответствие с существующим законодательством;
- возможность создания системы управления качеством в соответствии с международным стандартом ISO 9000:2000.

Внедрение в организации системы описания и управления БП, в первую очередь, подразумевает работу по реализации процессного подхода к управлению организацией.

Список литературы: 1. Хаммер М., Чампи Дж. Реинжиниринг корпорации: Манифест революции в бизнесе: Пер. с англ. СПб.: Издательство С.-Петербургского университета, 1997. 332 с. 2. Witten I.H., Bray Z., Mahoui M. and Teahan W.J. Text mining: a new frontier for lossless compression (1999) *Proc Data Compression Conference*, 198-207, IEEE Press, Los Alamitos, CA.

Поступила в редколлегию 11.04.2003

Левыкин Виктор Макарович, д-р техн. наук, профессор, заведующий кафедрой ИУС ХНУРЭ. Научные интересы: технологии разработки сложных информационных систем. Адрес: Украина, 6116, Харьков, пр. Ленина, 14, тел. 70-21-451.

Чалый Сергей Федорович, канд. техн. наук, доцент кафедры информатики ХНУРЭ. Научные интересы: проектирование информационных систем, моделирование бизнес-процессов. Адрес: Украина, 61166, Харьков, пр. Ленина, 14, тел. 70-21-419.