

УДК 658.5

ПРИМЕНЕНИЕ DCOR-МОДЕЛИ ПРИ МОДЕЛИРОВАНИИ БИЗНЕС-ПРОЦЕССОВ НА ОСНОВЕ СЕМАНТИЧЕСКОГО АНАЛИЗА ТЕКСТОВЫХ ДОКУМЕНТОВ

Зарайченко И.А., Галимулина Ф.Ф., Лубнина А.А.

*ФГБОУ ВО «Казанский национальный исследовательский технологический университет», Казань,
e-mail: irina-zar@mail.ru, 080502e_m@mail.ru, alsu1982@yandex.ru*

В статье предложена методика моделирования инновационных бизнес-процессов на основе семантического анализа текстовых документов в контексте нотаций DCOR-модели. Сущность методики заключается в последовательных этапах: семантический анализ концептов и анализ последовательности. Семантический анализ проводится в несколько этапов. Первый этап сводится к выявлению близости документов на основе расчета «весов» концептов. После этого рассчитывается семантическая близость концептов текущих регламентирующих документов и нотации DCOR-модели на основе алгоритма определения минимального расстояния между вершинами графа, определяющими близость концептов. Далее на основе сравнения последовательностей задач и операций проводится анализ последовательности выполнения операций. В результате реализации данной методики формируется концептуальная модель инновационного бизнес-процесса в контексте нотации DCOR, что является основой для последующей оптимизации структуры бизнес-процесса по следующим направлениям: исключение дублирования, объединение семантически схожих операций, применение лучших практик референтной модели. В отличие от традиционного подхода применение инструментов семантического анализа позволит повысить качество построенных моделей, а также сократить затраты человеческих ресурсов по их построению. Развитием данного метода является разработка инструментов автоматизированной оптимизации бизнес-процессов.

Ключевые слова: DCOR-модель, референтная модель, инновационный процесс, проектирование бизнес-процесса, семантический анализ, алгоритм, семантическая сеть, семантическая модель предметной области

APPLICATION OF THE DCOR MODEL IN MODELING BUSINESS PROCESSES BASED ON SEMANTIC ANALYSES

Zaraychenko I.A., Galimulina F.F., Lubnina A.A.

*Kazan National Research Technological University, Kazan,
e-mail: irina-zar@mail.ru, 080502e_m@mail.ru, alsu1982@yandex.ru*

The article proposes a technique for automated modeling of innovative business processes based on semantic analyses tools in the context of DCOR model notations. The essence of the methodology consists in sequential stages: semantic analysis of concepts, sequence analysis. Semantic analysis is carried out in several stages. The first stage is reduced to identifying the proximity of documents based on the calculation of the “weights” of concepts, then a semantic network of documents is formed and the degree of their proximity is assessed. After that, the semantic proximity of the concepts of the current regulatory documents and the notation of the DCOR-model is calculated based on the algorithm for determining the minimum distance between the graph vertices that determine the proximity of the concepts. Further, based on a comparison of the sequences of tasks and operations, the sequence of operations is analyzed. As a result of the implementation of this technique, a conceptual model of an innovative business process is formed in the context of the DCOR notation, which is the basis for the subsequent optimization of the structure of the business process in the following areas: elimination of duplication, unification of semantically similar operations, application of the best practices of the reference model. In contrast to the traditional approach, the use of semantic analyses tools will improve the quality of the constructed models, as well as reduce the cost of human resources for their construction. Development of this method is the development of tools for automated optimization of business processes.

Keywords: DCOR model, reference model, innovation process, business process design, semantic analysis, algorithm, semantic network, semantic domain model

Процессный подход в практику управления отечественными предприятиями вошел относительно недавно, в связи с этим острой проблемой стало построение качественной модели бизнес-процессов. Инструментом анализа и корректировки процессов в организации является бенчмаркинг на основе референтных моделей (SCOR, DCOR). Проблема применения этих моделей заключается в сложности построения схем процессов «как есть», так как имеет место склонность исполнителей субъектив-

но трактовать содержание, входы и выходы, последовательность операций и процедур в рамках реализации процесса. Такая ситуация осложняется многовариантностью бизнес-процессов в зависимости от конкретных условий их реализации, что в условиях динамичной внешней и внутренней среды становится краеугольным камнем построения эффективной системы бизнес-процессов, лежащей в основе их автоматизации. Проблема неопределенности внешней среды обостряется в рамках решения задачи

реализации инновационных процессов, так как данный тип бизнес-процессов характеризуется наиболее высоким уровнем неопределенности и стохастичности прогнозных моделей.

Данная проблема может быть решена на основе применения аналитических инструментов семантического анализа, основанных на применении цифровых инструментов анализа текстовых документов [1]. В основе данного подхода лежит исследование текстовых документов с помощью методов семантического анализа и последующее моделирование бизнес-процесса на его основе в рамках нотации DCOR-модели. Широкое распространение информационных технологий во всех сферах деятельности приводит к накоплению значительного объема данных не только в форме текстовых регламентирующих документов, но и в форме деловой переписки, данных ERP-систем, автоматизированных систем управления проектами и т.п. [2, 3]. Анализ такого рода информации на основе методов семантического анализа позволит минимизировать субъективное влияние исполнителей и исследователей на модель процесса и повысить ее качество.

Целью данного исследования является выявление особенностей применения инструментов семантического анализа для моделирования инновационных бизнес-процессов в рамках методологии DCOR-модели, а также разработка процедуры моделирования инновационного бизнес-процесса на основе инструментов семантического анализа. Отмеченная цель определяет перечень задач данной работы:

- изучить специфику применения инструментов семантического анализа регламентирующих документов, для построения модели бизнес-процесса разработки и коммерциализации инновационного продукта или технологии;
- исследовать количественные индикаторы семантического анализа регламентирующих документов;
- разработать процедуру моделирования инновационных бизнес-процессов, базирующуюся на методах семантического анализа и встраивания его результатов в методологию DCOR-модели.

Материалы и методы исследования

Перечень выбранных методов исследования обоснован поставленными задачами и характером предметной области. Были использованы:

- методология DCOR-модели, разработанная Международным Советом по логистике, включающая в себя рекомендации

по структуре и содержанию бизнес-процесса создания и коммерциализации инновационного продукта;

- методы семантического анализа текста, направленные на выявление концептов в документах, выявление семантической и смысловой связи концептов. Применение данного метода обосновано необходимостью анализа соответствия текущего состояния бизнес-процесса и адаптации его в рамках методологии DCOR. В частности, методика построения семантической сети и расчета веса концепта для визуализации и сравнения регламентных документов;

- метод графов в качестве основы для алгоритмов семантического анализа и расчета количественных показателей семантической связи концептов.

Синтез отмеченных методов в рамках процедуры моделирования позволяет разработать модель инновационного процесса, отражающую текущее состояние бизнес-процесса в нотации DCOR-модели.

Результаты исследования и их обсуждение

Методология DCOR-модели была предложена и актуализируется международным Советом по логистике и представляет собой логическое продолжение концепции SCOR-модели. Логика применения данных инструментов определяется используемым их в качестве стандартов и моделей совершенствования бизнес-процессов в цепях поставок. Если SCOR-модель определяет характер текущей деятельности организации, то DCOR-модель – Design Chain Operations Reference model (DCOR, Рекомендуемая модель операций в цепях проектирования) отражает особенности процессов создания и диффузии инноваций, их встраивания в структуру бизнес-процессов организации [4]. Объектами модели являются три различных направления инновационных процессов: Product Refresh (обновление продукта), New Product (новый продукт), New Technology (новая технология). В рамках данной модели выделено пять основных управленческих процессов: Plan (планирование), Research (исследование), Design (проектирование), Integrate (интеграция), Amend (совершенствование). Таким образом, данный инструмент представляет собой метод анализа и оптимизации бизнес-процессов организации, базирующийся на использовании «лучших практик» и сравнения их с текущим состоянием организации.

Применение DCOR-модели в качестве базового конструкта при моделировании бизнес-процессов организации обосновано ее универсальностью, а также актуальным

набором инструментов и лучших практик по выстраиванию инновационных процессов в организационных структурах мирового уровня. Вместе с этим применение DCOR-модели не является уникальным и новым направлением научных исследований, однако сложность представляет собой ее адаптация при наложении на текущие бизнес-процессы организации. Так, в работе С. Ou-Yang, I.H. Yin & Y.C. Juan [5] предлагается метод моделирования бизнес-процессов на основе семантического анализа и применения методологии качественного анализа полученных результатов на основе «обоснованной теории» (grounded theory). Методология семантического анализа для моделирования бизнес-процессов эффективно выстраивается на инструментах анализа Big data, так как включает в себя математический анализ закономерностей неколичественных данных: предполагается анализ степени близости терминологических определений основных этапов регламентированного бизнес-процесса и референтной модели. После этого существующий бизнес-процесс на основе анализа полученных количественных данных модернизируется на основе подходов референтной модели с целью исключения дублирования функций и избыточных процессов. Далее проводится оценка полученных данных на основе опросов исполнителей и качественного анализа полученных результатов. В целом данная методика представляет собой достаточно интересный подход, однако существенная сложность заключается в подходе к семантическому анализу, весьма чувствительному к качеству перевода терминологической базы описательной части процесса и DCOR-модели.

В связи с этой проблемой определяется необходимость адаптации методики семантического анализа в рамках исследуемого подхода для построения алгоритма автоматизации моделирования бизнес-процесса. В этой связи интересным представляется подход, представленный в работе А.В. Вицентия и др., который предлагает методологию машинного анализа текстовых документов на основе построения семантической модели предметной области (СМПО) [6], которая представляет собой «структуру взвешенных семантических отношений». СМПО позволяет реализовать процедуры извлечения и хранения множественного контекста употребленных в документах понятий, частично решая проблему совместимости новой информации с уже накопленными знаниями, а также выявить противоречия в семантических образах документов в случае, если новая информация

противоречит накопленной». Иными словами построение модели СМПО для бизнес-процесса разработки и коммерциализации нового продукта является базой для автоматизированного анализа и синтеза текущего состояния «как есть» и перспективного совершенствования на основе применения DCOR-модели. То есть фактически задачей построенной модели СМПО является формирование единообразной системы информации о бизнес-процессе, находящейся в одной «системе координат» с референтной моделью. Для построения семантической модели документа авторы предлагают использовать метод «"Semantic Network" – этот метод визуализирует семантическую модель документа в виде семантической сети. Семантическая сеть документа состоит из понятий документа и различных типов отношений между ними» [2]. Таким образом, задача анализа инновационного процесса сводится к построению семантической модели существующих регламентов, сравнение этой модели с документом DCOR-модели.

Вместе с этим актуальной остается задача математической интерпретации смысловых связей регламентирующих документов, анализ которых позволит реализовать задачи построения текущей модели бизнес-процесса. Один из подходов предполагает определение «веса» термина или семантического концепта, который определяется числом вхождений термина t в документе d , обозначаемым tf (term frequency) [7]. В ситуации наличия статистических данных об использовании терминов в наборе документов выполняется схема вычисления веса $tf-idf$, определяемая следующим образом:

$$idf_i = \log \frac{N}{df_i}, \quad (*)$$

где df – частота документа, определяемая как количество документов в коллекции, содержащих t , idf – обратная частота документа, N – общее количество документов в коллекции.

В рамках данного исследования была построена СМПО нотации DCOR-модели и регламента реализации бизнес-процесса «Разработка нового продукта» действующего производственного предприятия пищевой промышленности на основе метода «Top 100 words» (табл. 1).

Как видно из табл. 1, данный вид анализа позволил выявить высокую степень различий в семантике двух документов. В результатах анализа среди семантического ядра документа выявлены только два концепта, схожие в двух документах.

Для оценки семантической связи документов целесообразно исследовать регламентные документы и актуальную версию DCOR-модели на основе алгоритма, представленного в работе М.И. Варламова и А.В. Коршунова, который лежит в основе информационно-аналитической системы Текстера [8]. Согласно данному алгоритму автор оценивает степень смысловой близости концептов посредством анализа графов ссылок Википедии и расчета величины кратчайшего пути [8]:

$$\min\{\delta_{uv} + \delta_{vw} | w, \delta_{uv} \in L(u), w, \delta_{vw} \in L(v)\},$$

где $L(u)$, $L(v)$ – метки вершин u и v (концепты – ссылки), δ_{uv} – расстояние между u и w , и расстояние между w и v соответственно.

На основе данного алгоритма оценивается соответствие текущих процессов и процедур в структуре бизнес-процесса

с аналогичными процессами, описанными в методологии DCOR. При этом важной особенностью является расчет двух показателей семантической близости: показатель действия элемента процесса и показатель объекта элемента процесса. Например, в регламентирующем документе выделен процесс «формирования идеи», данный процесс включает действие, обозначенное термином «формирование» и объектом «идеи», аналогично для процесса «R2.1 Receive & Validate Request» (получение и идентификация запроса) в DCOR-модели можно выделить блок «действия» (получение и идентификация) и блок «объекта» (запрос), поэтому оценка семантической связи предполагает две оценки: оценка связи «действия» и «объекта». В табл. 2 представлен фрагмент структуры использованных в исследовании концептов.

Таблица 1

Результаты количественного семантического анализа по методу «Top 100 words» (фрагмент)

Регламент бизнес-процесса	TF-IDF	Описание DCOR-модели (APICS)	TF-IDF
Проекта	124	Процесс	140
Продукта	91	Элемент	129
Руководитель	81	Модель	128
Процесс	79	Уровень	125
нового продукта	78	Цепочки	118
план	77	Планирование	117
руководитель проекта	76	элемент процесса	115
этапе	74	Метрика	114
затраг	67	связанной	113
рынок	67	показатели	112
документа	65	дизайн	111
prd	64	исследование	111
вывод	64	проектирование	111
проектирование	63	производительности	111
продукта на рынок	63	интеграция	110
проекта команда	62	поставок	110
работ	62		

Таблица 2

Структура концептов, использованных в исследовании

	Действие	Объект
Регламент		
1.1. Формирование идеи	Формирование	Идея
1.2. Работы по регистрации знака	Регистрация	Знак
2. Формирование и утверждение проекта	Формирование, утверждение	Проект
3. Формирование плана проекта	Формирование, план	Проект
4. Разработка и освоение производства нового продукта	Разработка, производство	Продукт
5.1. Подготовка запуска вывода продукта на рынок	Подготовка, вывод	Рынок
DCOR-модель		
R2.1. Receive & Validate Request	Receive, validate	Request
R2.2. Schedule Research Activities	Schedule, Research	Activities
R2.5. Transfer Findings /Materials	Transfer	Materials
R2.6. Authorize Supplier Payment	Authorize	Payment

Таблица 3

Результаты расчета показателей семантической связи для некоторых концептов бизнес-процессов

	D2.1 Получение, подтверждение и декомпозиция	D2.2 Планирование производственных мероприятий	ED.2 Управление показателями производства	ED.3 Управление информацией о производстве	ED.5 Управление производственными мощностями	ED.8 Мониторинг соответствия новым показателям	ED.9 Управление интеллектуальной собственностью
1.1. Формирование идеи	(0,62; 0,53)	(0; 0,01)	(0,08; 0)	(0,12; 0)	(0,0; 24)	(0,35; 0)	(0,06; 0,02)
1.2. Работы по регистрации знака	(0,12; 0)	(0,08; 0,13)	(0,12; 0,22)	(0,11; 0)	(0; 0,03)	(0; 0,01)	(0,51; 0,67)
2. Формирование и утверждение проекта	(0; 0,01)	(0,73; 0,44)	(0; 0,01)	(0,07; 0,04)	(0,12; 0)	(0,06; 0,02)	(0; 0,14)
3. Формирование плана проекта	(0; 0)	(0,72; 0,58)	(0,11; 0,05)	(0; 0)	(0,14; 0,08)	(0; 0)	(0,05; 0,09)
4. Разработка и освоение производства нового продукта	(0; 0,07)	(0; 0)	(0,07; 0,17)	(0; 0,08)	(0,61; 0,3)	(0,12; 0)	(0; 0,01)
5.1. Подготовка запуска вывода продукта на рынок	(0,11; 0,05)	(0; 0,03)	(0,71; 0,48)	(0; 0,04)	(0,06; 0,02)	(0,11; 0)	(0; 0)
5.2. Обеспечение достижения целевых рыночных задач проекта	(0; 0,01)	(0; 0)	(0,12; 0)	(0,51; 0,43)	(0,08; 0,13)	(0,61; 0,72)	(0,05; 0,1)
6. Оценка и контроль финансовых показателей проекта	(0; 0)	(0,21; 0,11)	(0,48; 0,52)	(0; 0)	(0,07; 0,04)	(0,06; 0,2)	(0; 0,02)

Результаты расчета показателей семантической связи для некоторых концептов бизнес-процессов представлены в табл. 3. Цветом в таблице выделены значения семантической близости, имеющие значимые величины. Таким образом, на основе семантического анализа было выявлено соответствие бизнес-процессов регламента предприятия и бизнес-процессов DCOR-модели. На основе выявленной связи построим схему реализации бизнес-процесса в нотации DCOR-модели (рисунок).

– применения описанных «лучших практик» для совершенствования эффективности бизнес-процесса.

Заключение

Таким образом, можно систематизировать процедуру анализа процесса разработки и коммерциализации инновационного продукта как последовательность следующих этапов:

– анализ структуры и весов терминов *idf* с целью определения соответствия общей

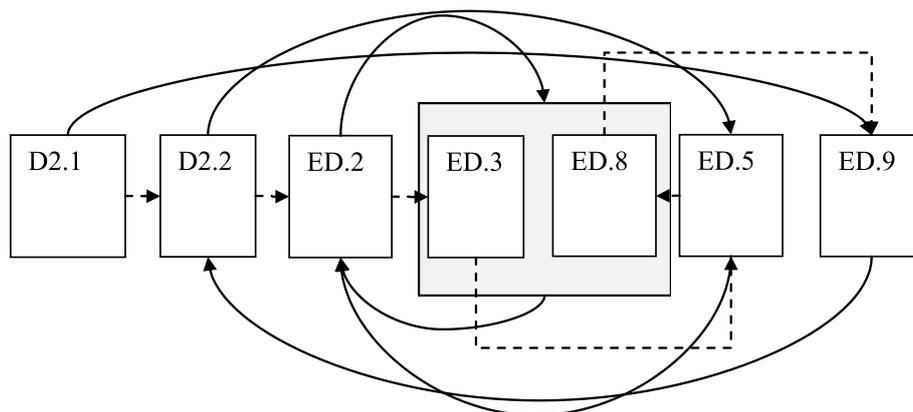


Схема последовательности выполнения операций «как есть».

Примечание. Пунктирной линией обозначена логика DCOR-модели, сплошной линией – фактическая логика реализации процесса в исследуемой организации

На основе результатов сравнительного моделирования текущего бизнес-процесса в нотации DCOR можно сделать следующие выводы:

- процессы ED3, ED8 в текущем регламенте представляют собой единый процесс;
- последовательность реализации элементов бизнес-процесса не соответствует логике DCOR-модели;
- на основе выявленного соответствия можно применять сформулированные в DCOR-модели лучшие в практике в рамках аналогичных бизнес-процессов.

На данном этапе сформирована модель текущего состояния бизнес-процесса создания нового продукта в контексте нотации DCOR-модели. Данная схема является объективным описанием модели бизнес-процесса «как есть» и является основой для дальнейшей оптимизации. На основе данной модели можно решить следующие задачи оптимизации бизнес-процесса:

- выявление дублирующих задач, функций, процессов (отражает цикличность процессов);
- необходимость объединения процессов на основе выявления семантической связи описания содержания процесса;

структуры и содержания бизнес-процессов с методологией DCOR. На данном этапе сравнивается перечень выявленных концептов, а также сравнение полученных «весов» с аналогичными значениями текущей версии DCOR-модели, разработанной Международным Советом по логистике;

– далее на основе метода «semantic network» строится модель регламентирующих документов, отражающая схематическую модель инновационного процесса организации и сравнения ее структуры с семантической сетью нотации DCOR-модели;

– этап количественного автоматизированного семантического анализа на основе алгоритма определения кратчайшего пути графа, что позволяет выявить наиболее семантически близкие описания бизнес-процессов и процедур в регламентирующих документах и нотации DCOR-модели;

– построения модели последовательности выполнения процедур и процессов в текущем состоянии в нотации DCOR.

Развитием данного исследования является разработка методики автоматизированной оптимизации моделей бизнес-процессов посредством применения инструментов

машинного обучения. Безусловно, автоматизация систем моделирования бизнес-процессов позволяет существенно развивать управленческие технологии, в том числе за счет снижения трудоемкости аналитической работы по обработке неструктурированных данных о бизнес-процессах. В отличие от традиционного подхода управление на основе количественных данных позволяет решать новые управленческие задачи, в частности задачи моделирования бизнес-процессов на основе анализа текстовых документов. Вместе с этим очевидна необходимость развития методов семантического анализа в контексте отмеченных задач, так как текущие подходы в большей степени ориентированы на решение маркетинговых задач. Современные автоматизированные системы управления в наибольшей степени ориентированы на решение задач контроля и планирования на основе количественных данных, в то время как значительная часть управленческих решений в области построения и моделирования бизнес-процессов остается прерогативой человека. Вместе с этим в условиях динамичных управленческих задач возникает острая необходимость автоматизации решения такого рода задач, что обусловлено необходимостью, с одной стороны, оптимизировать затраты управленческих ресурсов, а с другой, устранить вероятность субъективности и неэффективности принимаемых в данной области управленческих решений.

Таким образом, предлагаемая методика позволяет применять методы и инструменты семантического анализа текстовых

документов для построения модели бизнес-процесса в нотации DCOR-модели, что предполагает последующую модернизацию процесса на основе лучших практик.

Исследование выполнено в рамках гранта Президента РФ по государственной поддержке ведущих научных школ РФ № НШ-2600.2020.6

Список литературы

1. Шинкевич А.И., Барсегян Н.В. Пути повышения эффективности организации производственных процессов на нефтехимических предприятиях за счет применения систем автоматизации // Русский инженер. 2019. № 4. С. 48–51.
2. Lubnina A.A., Shinkevich M.V., Yalunina E.N., Gaydamashko I.V., Savderova A.F., Komissarova M.A. Innovative strategy for improving the efficiency of industrial enterprises management. *Espacios*. 2018. Т. 39. № 9. P. 25.
3. Сучков М.А., Галимулина Ф.Ф. Принципы управления криптоданными в рамках инновационного развития информационной среды предприятия // Наука и бизнес: пути развития. 2020. № 5 (107). С. 152–154.
4. Сергеев В.И. DCOR моделирование как развитие стандарта SCOR-модели // Логистика и управление цепями поставок. 2008. № 4 (27). С. 8–13.
5. Ou-Yang C., Yin I.H., Juan Y.C. A DCOR-based NPD process redesign approach with grounded theory evaluation. *South African Journal of Industrial Engineering*. 2019. Vol. 30 (2). P. 45–60.
6. Вицентий А.В., Шишаев М.Г., Диковицкий В.В. Опыт использования визуализации семантических моделей документов для экспресс-контент-анализа (на примере нормативно-правовых актов, регламентирующих развитие пространственно-распределенных систем) // Труды Кольского научного центра РАН. 2018. № 10–9 (9). С. 99–111.
7. Salton G., McGill M.J. *Introduction to modern information retrieval*. McGraw-Hill, 1983.
8. Варламов М.И., Коршунов А.В. Расчет семантической близости концептов на основе кратчайших путей в графе ссылок Википедии // Машинное обучение и анализ данных. 2014. Т. 1. № 8. С. 1107–1125.