

ОТРАСЛЕВАЯ И РЕГИОНАЛЬНАЯ ЭКОНОМИКА

Р. Д. Гимранов¹

ПАО «Сургутнефтегаз» (Сургут, Россия)

С. А. Тищенко²

МГУ имени М. В. Ломоносова (Москва, Россия)

М. А. Шахмурадян³

МГУ имени М. В. Ломоносова (Москва, Россия)

П. О. Вакорин⁴

МГУ имени М. В. Ломоносова (Москва, Россия)

А. А. Выслоух⁵

МГУ имени М. В. Ломоносова (Москва, Россия)

М. О. Коматовский⁶

МГУ имени М. В. Ломоносова (Москва, Россия)

ГРАФ ЦИТИРОВАНИЙ КАК ИНСТРУМЕНТ МЕТОДОЛОГИИ ИССЛЕДОВАНИЯ НАУЧНОЙ ЛИТЕРАТУРЫ ПО ОНТОЛОГИИ БИЗНЕС-ПРОЦЕССОВ ПРЕДПРИЯТИЯ

Цифровизация бизнес-процессов предприятия является неотъемлемой частью современной конкурентоспособной политики коммерческой организации. Использование результатов качественного и количественного анализов существующих научных подходов в области определения и развития онтологий, опосредующих производство товаров/услуг, может положительно отразиться на принятии стратегических решений ввиду сокращения времени на обзор и исследование однотипных идей и методов. Статья посвящена использованию элементов теории графов для анализа множества концептуальных школ и подходов к онтологии бизнес-процессов. Авторами

¹ Гимранов Ринат Дамирович, начальник управления информационных технологий, ПАО «Сургутнефтегаз»; e-mail: gimranov_rd@mail.ru

² Тищенко Сергей Александрович, к.физ.-мат.н., доцент экономического факультета МГУ имени М. В. Ломоносова; e-mail: tichtch@mail.ru

³ Шахмурадян Михаил Андреевич, аспирант экономического факультета МГУ имени М. В. Ломоносова; e-mail: mshakhmuradyan@econ.msu.ru

⁴ Вакорин Павел Олегович, студент бакалавриата экономического факультета МГУ имени М. В. Ломоносова

⁵ Выслоух Андрей Андреевич, студент бакалавриата экономического факультета МГУ имени М. В. Ломоносова

⁶ Коматовский Максим Олегович, студент бакалавриата экономического факультета МГУ имени М. В. Ломоносова

было проанализировано 216 научных источников по данной тематике. В результате исследования был построен граф цитирований, систематизирующий отобранные по описанной выше методологии работы авторов изучаемой сферы. Данный инструмент методологии анализа может быть масштабирован и использован заинтересованными лицами для более качественного анализа литературы по онтологии бизнес-процессов предприятия.

Ключевые слова: бизнес-процессы, граф цитирований, онтология бизнес-процессов, методология исследования.

Цитировать статью: Гимранов Р. Д., Тищенко С. А., Шахмурадян М. А., Вакорин П. О., Выслоух А. А., Коматовский М. О. Граф цитирований как инструмент методологии исследования научной литературы по онтологии бизнес-процессов предприятия // Вестник Московского университета. Серия 6. Экономика. — 2019. — № 6. — С. 99–110.

R. D. Gimranov

PJSC Surgutneftegaz (Surgut, Russia)

S. A. Tishchenko

Lomonosov Moscow State University (Moscow, Russia)

M. A. Shakhmuradyan

Lomonosov Moscow State University (Moscow, Russia)

P. O. Vakorin

Lomonosov Moscow State University (Moscow, Russia)

A. A. Vysloukh

Lomonosov Moscow State University (Moscow, Russia)

M. O. Komatovsky

Lomonosov Moscow State University (Moscow, Russia)

CITATION GRAPH AS AN INSTRUMENT OF THE RESEARCH METHODOLOGY OF SCIENTIFIC LITERATURE ON ONTOLOGY OF ENTERPRISE BUSINESS PROCESSES

The digitalization of enterprise's business processes is an integral part of the modern competitive policy of a commercial organization. The use of the results of qualitative and quantitative analyzes of existing scientific approaches in the field of identifying and developing ontologies that mediate the production of goods / services can positively affect the adoption of strategic decisions due to the reduction of time for review and research of the same ideas and methods. The article is devoted to the use of graph theory elements for the analysis of many conceptual schools and approaches to the ontology of business processes. The authors analyzed 216 scientific sources on this topic. As a result of the study, a citation graph was constructed that systematized the authors of the study area selected according to the methodology described above. This tool of the analysis methodology can be scaled and used

by interested parties for a better analysis of the literature on the ontology of enterprise business processes.

Key words: business processes, citation graph, ontology of business processes, research methodology.

To cite this document: *Gimranov R. D., Tishchenko S. A., Shakhmuradyan M. A., Vakorin P. O., Vysloukh A. A., Komatovsky M. O.* (2019). Citation graph as an instrument of the research methodology of scientific literature on ontology of enterprise business processes. *Moscow University Economic Bulletin*, (6), 99–110.

Введение

В условиях экспоненциального уровня развития информации эффективность деятельности современного предприятия опосредуется рациональностью его бизнес-процессов. Для рентабельного развития на рынке в настоящее время организациям приходится искать способы ведения бизнеса, максимизирующие показатели их конкурентноспособности, а также минимизирующие затраты на данный процесс. Одним из таких способов является внедрение онтологии в бизнес-процессы предприятия. Термин «*онтология*» имеет значительное число вариаций определения, но в ходе изучения данной тематики авторы условились остановиться на одном, наиболее подходящем для проблематики данной работы, а именно дефиниции, предложенной профессором Технологического института Карлсруэ Руди Студером в 1998 г.: «Онтология — это формальная спецификация разделяемой концептуальной модели». Иными словами, онтология бизнес-процессов представляет собой структурную спецификацию процессов, их формализованное представление, которое включает словарь (или имена) указателей на термины данных процессов и их логические выражения, которые описывают, как они соотносятся друг с другом [Гаврилова, 2002, р. 845]. Авторы данной статьи решили провести обзор некоторых ключевых работ авторов, специализирующихся на изучении онтологии бизнес-процессов (*Business Process Ontology*) и онтологии предприятия (*Enterprise Ontology*), а также смежных с данной тематикой сфер, таких как семантические и прагматические сети (*Semantic & Pragmatic Web*), связанные с ними модели (*RDF-триплеты*), различные языки онтологий для описания семантических сетей (*OWL*), а также инженерия предприятий (*Enterprise Engineering*). В настоящей работе описана основная методология вышеуказанного исследования, которая позволила авторам систематизировать процесс изучения литературы, а также выделить определенные реперные точки (основополагающие работы исследуемых тематик). Статья состоит из введения, описания основной используемой терминологии из теории графов, информации о сборе данных, методологии построения графа, анализа тенденции развития знания об онтологии бизнес-процессов и заключения.

Основная информация из теории графов

Перед описанием результатов исследования авторов необходимо ввести ключевые понятия из теории графов. *Граф* — это пара множеств (V, E) , где V (англ. vertex — вершина) — множество вершин, а E (англ. edge — ребро) — множество упорядоченных пар вершин, обозначающих связи между ними (множество ребер).

Кратность (разряд) (англ. degree) вершины — количество вершин, связанных с данной. Стоит отметить, что для ориентированных графов вводятся понятия *in-degree* и *out-degree*, равные количеству входящих и исходящих дуг соответственно.

Петля (loop) — ребро, начало и конец которого совпадают. Вершина, подобным образом, связывается сама с собой, повышая свою кратность. Стоит отметить, что большинство авторов не считают петли разновидностью *циклов*.

Связный граф — это граф, для любой пары вершин которого существует путь, их соединяющий.

Мера центральности (центральность) — функция на множестве вершин графа, упорядочивающая вершины в соответствии с критерием центральности (например, по количеству связей). Служит для выявления наиболее «важных» вершин.

Престиж Катца—Боначича — мера центральности, критерием которой выступают собственные значения матрицы смежности сети. Матрица смежности строится согласно следующему принципу:

$$D_{i,j} = \begin{cases} 1, & \text{если } (i, j) \in E \\ 0, & \text{если } (i, j) \notin E \end{cases}$$

Онтология — это формальная спецификация разделяемой концептуальной модели.

Сбор данных

Изначально коллективом авторов были проанализированы три фундаментальных научных исследования по тематике, авторами которых являлись Александр Остервальдер, Ян Дитц и Джон Захман. Дальнейшие работы определялись либо прямым следованием рекомендациям упомянутых авторов, либо с использованием поиска по ключевым словам в системах *Google Scholar*, *Springer* и прочих информационных агрегаторах. Стоит отметить, что рассматривались *все* источники, имевшие отношение к онтологии бизнес-процессов, вне зависимости от времени написания. Это делает выборку изученных работ приблизительно независимой от хронологии публикации, т.е. можно без ограничения общности предположить, что в любой период времени нами была найдена определенная

фиксированная доля всех публикаций. Вершинами графа, для построения которого и производился сбор данных, являются авторы, а ребрами — цитирования. Автор включался в выборку только в том случае, если авторы совместно приходили к выводу о том, что его работы привносят в изучаемую область некоторую новизну, а общие идеи связаны с онтологией бизнес-процессов, бизнес-моделями и реинжинирингом предприятий.

Методология построения графа

Далее отражены основные характеристики и методологические особенности построения графа цитирований.

1) *Граф рассматривается и как ориентированный* (ребрам которого присвоено направление), *и как неориентированный*, в зависимости от того, какая конфигурация требуется для применения конкретной аналитической методики (так, например, для подсчета мер центральности Катца—Боначича имеет смысл рассматривать лишь ориентированный, а для подсчета мер кластеризации Ньюмана — ассоциированный с ним неориентированный прототип).

2) *Все ребра имеют равный вес*. Подавляющая часть литературы носит методологический характер, удалось найти лишь 5–6 источников, описывающих именно прикладной аспект использования онтологий (онтология технических процессов электростанции, онтология некоторого класса химических элементов, использующихся на конкретном заводе, и т.д.).

3) *Исключены петли и множественные связи*. Петли (цитирование автором самого себя) происходили лишь в случае цитирования группы соавторов, в составе которых находился цитирующий. Эти весьма распространенные случаи лишь завышали и без того высокие показатели центральности наиболее популярных исследователей, не добавляя при этом никакой информации. Множественные связи исключены по идентичным причинам.

4) *Группы соавторов воспринимались как совокупности участников*. Например, если группа {А, В} процитировала группу {С, D, E}, в графе добавлялось максимум 6 ребер и 5 вершин, если никакие из этих ребер и вершин ранее не встречались.

5) *Размер вершины в графе зависел от in-degree этой вершины*, т.е. чем идеи автора популярнее среди коллег, тем он «больше». В силу особой популярности некоторых авторов пришлось произвести разумное логарифмическое масштабирование.

б) *Авторы кодировались порядковым номером*, вершины нумеровались¹.

¹ Для получения декодера авторов научных работ — пишите авторам данной работы на электронную почту.

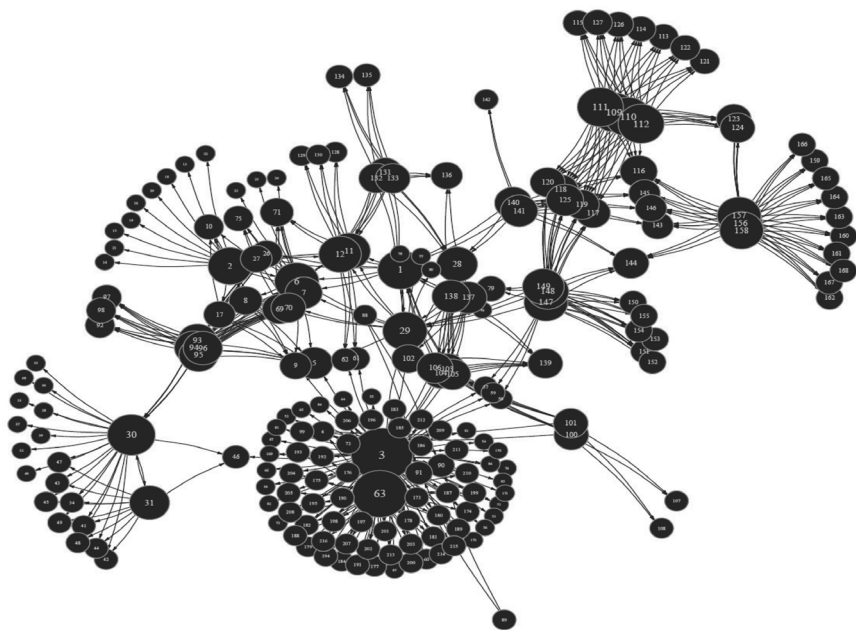


Рис. 1. Граф цитирований
Источник: составлено авторами.

Разбиение на группы влияния

Всего в построенном графе насчитывается 216 вершин. Количество связей с учетом петель и повторных ребер равно 734. Без учета петель и повторов — 559. Размер вершин зависит от степени следующим образом:

$$\text{Size} = 3 + 4\log_2(\text{deg} + 1).$$

Для выявления наиболее влиятельных направлений использовалась мера центральности Катца [Katz, 1953, p.39] и Бонаичича [Bonacich, 1987], описанная в работе Мэтью Джексона [Jackson, 2008]. Поскольку эта мера на неориентированной сети совпадает с обычной степенной (*англ.* degree centrality), можно просто покрыть всю сеть (или ее большую часть) минимальным набором окрестностей вершин с наибольшими степенями. Из чего следует результат, отраженный на рис. 2. При наблюдении графа очевидны некоторые недостатки: во-первых, это вложенные окрестности, например, у вершин 3 и 63, а во-вторых, это непокрытые области (пучок с началом в 156–158 и др.).

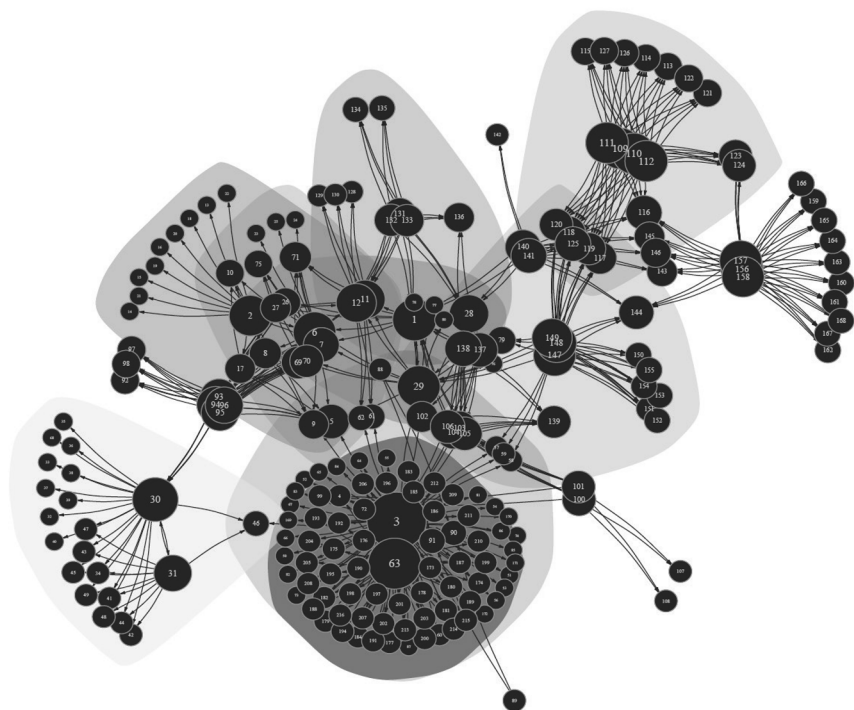


Рис. 2. Покрывание графа окрестностями 10 вершин с наибольшими степенями

Источник: составлено авторами.

На этом этапе список наиболее влиятельных авторов был сокращен вручную до девяти фамилий. Полученные окрестности полностью покрыли сеть, а исследователи, вокруг которых эти окрестности строились, отразили *девять* основных направлений мысли (см. рис. 3).

Изучение каждой области состоит из трех частей: общая характеристика, подход «лидера мнений», подходы предшественников/последователей лидера мнений. Отдельное внимание уделяется тем, кто попал в «пограничную зону», т.е. вошел в несколько областей, а также тем, кто не вошел ни в какую из них (последние, вероятно, либо исследуют область, в которой академическое сообщество не видит потенциала, либо занялись новым направлением). Также заметим, что одна из групп, полностью покрытая соседними, не была исключена из анализа. Эта группа — окрестность А. Остервальдера, отличающаяся особым интересом к агрегированию существующих работ, — имеет прямое пересечение с пятью из восьми остальных групп.

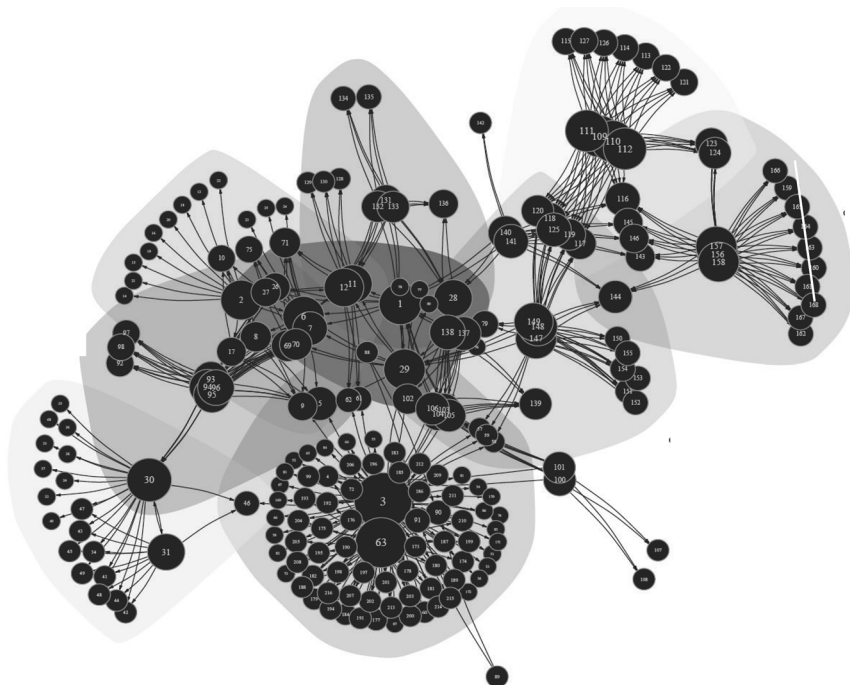


Рис. 3. Итоговое разбиение графа на области
 Источник: составлено авторами.

Анализ тенденций развития области знания и специфика процедуры сбора информации

Пытаясь ответить на вопрос об истории развития области изучения онтологий, текущего состояния дел и выявления потенциальных сценариев будущих исследований был построен график накопленного количества публикаций в период с 1960 по 2015 г. (см. рис. 4).

Глядя на иллюстрацию, нетрудно заметить, что основное развитие области происходило в период с 1990 по 2005 г., после чего наблюдается существенное снижение темпов роста количества публикаций (выпуклость функции сменяется вогнутостью). Объяснить подобное «снижение интереса» можно совокупностью трех причин.

1) Онтологические исследования, отвечая на запросы мобильности, переместились в интернет-источники (например, один из передовых прикладных исследователей в области онтологии бизнес-процессов Стивен Вольфрам в течение весьма длительного времени ведет личный блог, посвященный структурированию различных областей научного знания).

2) Многие онтологические исследования производятся локально, например в рамках конкретного предприятия, поскольку не всегда возможно применение общих методов систематизации в каждом конкретном случае. Приходится некоторым образом подстраивать исследование под специфику отрасли, предприятия, отдела и т.д. Довольно естественно предположить, что подобного рода работы не попадают в научные журналы, поскольку содержат некоторую корпоративную информацию, не подлежащую широкому распространению.

3) Подобная форма кривой частично может быть объяснена алгоритмом поиска статей. Представим, что существует некоторая истинная сеть авторов, связанных ребрами публикаций. Распространение наших знаний об этой сети из начальных точек (Остервальдер, Дитц и Захман) подобно некоторому диффузионному процессу (например, подобным образом распространяется инфекционное заболевание в сети, описывающей население, в случае, когда болезнь передается по ребрам графа каждому из соседей зараженного агента с некоторой постоянной вероятностью). Такого рода распространение может быть описано, например, диффузионной моделью Ф. Басса [Bass, 1969, p. 215].

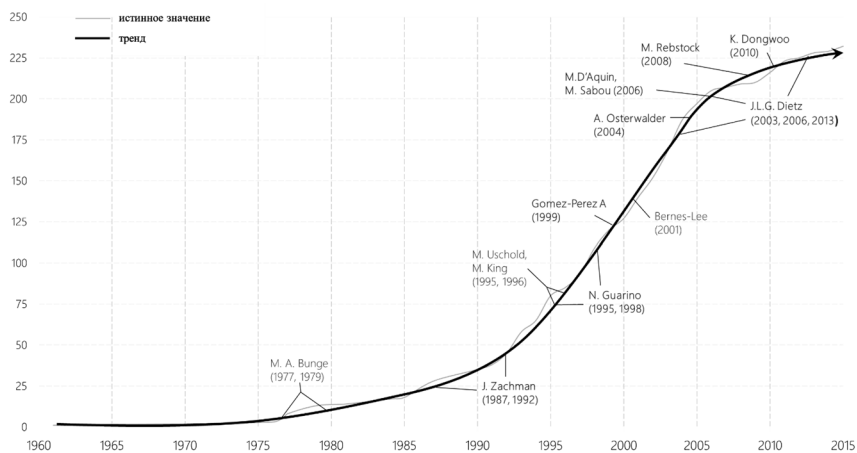


Рис. 4. График накопленного количества научных работ на темы, связанные со сферой онтологии предприятия
Источник: составлено авторами.

Формальные показатели

Число замкнутых треугольных структур составило 248. Число незамкнутых треугольных структур — 17 492. Из этого следует, что общий показатель кластеризации (*overall clustering*) в смысле Ньюмана [Newman,

Girvan, 2004] (отношение количества замкнутых треугольных структур к количеству незамкнутых треугольных структур) равен:

$$CI = 0,014177.$$

Обычно чем выше данный показатель, тем плотнее сеть.

Кластером в теории графов называют подграф, по плотности близкий к полному. Высокая степень кластеризации графа цитирований возможна лишь в рамках некоторой узкой области, где крайне ограниченный круг специалистов на постоянной основе занимаются взаимным цитированием. В случае развития нескольких идейных направлений общий показатель кластеризации редко поднимается выше 5%. Среднее число цитирований в расчете на автора составило приблизительно 2,52. Диаметр ненаправленной сети составил 8. Диаметр наибольшего связного ориентированного подграфа — 4. Это достаточно существенное значение, поскольку, для сравнения, приблизительный диаметр открытой части сети интернет оценивается числами в окрестности 7. Вместе эти показатели говорят о разнонаправленности движения мысли в научной области по изучению онтологии бизнес-процессов предприятия.

Заключение

В заключение стоит отметить, что в ходе проведения анализа, изложенного в данной работе, был выявлен эффективно визуализирующий инструмент для дальнейшего анализа академической литературы, основными темами которой являются онтологии бизнес-процессов, а также смежных с данной тематикой сфер исследования. В ходе работы был построен граф, систематизирующий отобранные по описанной выше методологии работы авторов исследуемой сферы. Были выявлены реперные точки — ключевые деятели своих областей интереса с соответствующими им областями — исследователями, прямо или косвенно с ними связанными. Параллельно с количественным анализом авторами также проводился и качественный анализ по шести параметрам практического применения основных направлений научных работ по теме онтологии бизнес-процессов организации, результаты которого по причине объема будут отражены в другой научной статье.

Использование графа цитирований, по мнению авторов данной статьи, является действенным методом проведения исследования (среднее число цитирований в расчете на автора составило 2,52, диаметр наибольшего связного ориентированного подграфа 4, общий показатель кластеризации в смысле Ньюмана 0,014177). Данный метод может быть применен исследователями для более качественного анализа литературы по онтологии бизнес-процессов предприятия. В дальнейшем авторы планируют продолжать научные изыскания в области применения инструменталь-

ных методов к анализу научной литературы об онтологии бизнес-процессов предприятия.

Список литературы

1. *Гаврилова Т. А.* Онтологический инжиниринг // Труды конференции «КИИ-2002». — М., 2002. — С. 845–853.
2. *Любкин А. А.* Введение в теорию графов: учебно-методическое пособие. — Изд-во МГУ, 1975. — 135 с.
3. *Bass F. M.* A New Product Growth for Model Consumer Durables // Management Science. Theory Series. — January 1969. — Vol. 15. — No. 5. — P. 215–227.
4. *Berners-Lee T., Hendler J., Lasilla O.* The Semantic Web, Scientific American, May 2001.
5. *Bonacich P.* Power and centrality: a family of measures // American Journal of Sociology. — 1987. — 92. — 1170–1182.
6. *Date C. J.* An Introduction to Database Systems // Addison-Wesley Publishing Co., Reading, MA, 1981.
7. *Dongwoo K., Jeongsoo L., Sungchul C., Kwangsoo K.* An ontology-based Enterprise Architecture, Expert System with Application. — 2010. — 37. — P. 1456–1464.
8. *Gruber T.* Towards principles for the design of ontologies used for knowledge sharing, 1995.
9. *Gruninger M., Fox M. S.* An Activity Ontology for Enterprise Modelling. — Department of Industrial Engineering, University of Toronto, 1994.
10. *Guarino N.* Formal Ontologies and Information Systems / in Guarino N. (ed), Proc. of FOIS98, IOS Press, 1998. — P. 3–15.
11. *Jackson M. O.* Social and Economic Networks. Princeton University Press, 2008.
12. *Jensen K.* Colored Petri Nets. Basic Concepts / Analysis Methods and Practical Use. Volume 1, Basic Concepts. Monographs in Theoretical Computer Science, Springer Verlag, 1997.
13. *Lopez M. F.* Building a Chemical Ontology Using Methontology and the Ontology Design approach, 1999.
14. *Lopez M. F., Corcho O.* Ontological Engineering, 2004.
15. *Malone T. W., Yates J., and Benjamin R. I.* Electronic Markets and Electronic Hierarchies, Communications of the ACM 30, No. 6, 484497, June 1987.
16. *Newman M. E., Girvan M.* Finding and evaluating community structure in networks, Phys. Rev. E 69, 2004.
17. *Pazos S. J.* Building Ontologies at the Knowledge Level using the Ontology Design Environment, 1998.
18. *Pazos S. J.* Enterprise Modelling, 1998.
19. *Sowa J. F.* Semantic Networks // Encyclopedia of Artificial Intelligence, Second Edition, S. C. Shapiro, Editor, Wiley, New York, 1992. — P. 1493–1511.
20. *Spewak S. H., Steven C. H.* Enterprise Architecture Planning: Developing a Blueprint for Data Application and Technology. — NY: John Wiley Sons Inc, 1992.
21. *Uschold M.* The enterprise ontology, 1997.
22. *Uschold M., Jasper R.* A Framework for Understanding and Classifying Ontology Applications / Proceedings of the UCAI-99 workshop on Ontologies and Problem-Solving Methods (KRRS), 1999.

23. *Uschold M., King M., Moralee S., Zorgios Y.* The Enterprise Ontology, AIAI, The University of Edinburgh, 1995.

The List of References in Cyrillic Transliterated into Roman Alphabet

1. *Gavrilova T. A.* Ontologicheskij inzhiniring // Trudy konferencii «KII-2002». — M., 2002. — S. 845—853.
2. *Ljubkin A. A.* Vvedenie v teoriju grafov Uchebno-metodicheskoe posobie. — Izd-vo MGU, 1975. — 135 s.